COMILLAS
UNIVERSIDAD PONTIFICIA
ICAI

GRADO EN INGENIERÍA EN TECNOLOGÍAS DE TELECOMUNICACIÓN

TRABAJO FIN DE GRADO

CHARACTERIZING TRADEOFFS IN SENSOR AVAILABILITY AND FAULT DETECTION EFFICACY

Autor: Ignacio Sanz Barro

Director: Gabe Fierro

Madrid

June 2025

Declaro, bajo mi responsabilidad, que el Proyecto presentado con el título

CHARACTERIZING TRADEOFFS IN SENSOR AVAILABILITY AND FAULT DETECTION EFFICACY

en la ETS de Ingeniería - ICAI de la Universidad Pontificia Comillas en el

curso académico 2024/25 es de mi autoría, original e inédito y

no ha sido presentado con anterioridad a otros efectos.

El Proyecto no es plagio de otro, ni total ni parcialmente y la información que ha sido

tomada de otros documentos está debidamente referenciada.

Fdo.: Ignacio Sanz Barro Fecha: 26/ 06/ 2025

Autorizada la entrega del proyecto

EL DIRECTOR DEL PROYECTO

Fdo.: Gabe Fierro Fecha: 26/ 06/ 2025

COMILLAS
UNIVERSIDAD PONTIFICIA
ICAI

GRADO EN INGENIERÍA EN TECNOLOGÍAS DE TELECOMUNICACIÓN

TRABAJO FIN DE GRADO

CHARACTERIZING TRADEOFFS IN SENSOR AVAILABILITY AND FAULT DETECTION EFFICACY

Autor: Ignacio Sanz Barro

Director: Gabe Fierro

Madrid

June 2025

Acknowledgments

I want to transfer my most sincere appreciation to Colorado School of Mines for its warm and welcoming atmosphere since my arrival. This last year has been a deeply enriching journey, academically, personally and professionally.

I want to give special recognition to Professor Gabe Fierro. I appreciate his dedication and generosity a lot, and it has contributed significantly to my academic accomplishments. I am grateful to him for his guidance throughout this project, as he has been helping and guiding me throughout this project.

CARACTERIZACIÓN DEL EQUILIBRIO ENTRE LA DISPONIBILIDAD DE SENSORES Y EFICACIA EN LA DETECCIÓN DE FALLOS

Autor: Sanz Barro, Ignacio.
Director: Fierro, Gabe.
Entidad Colaboradora: Colorado School of Mines

RESUMEN DEL PROYECTO

En este trabajo se estudian las ventajas entre la disponibilidad de sensores y la eficacia de la detección de fallos en sistemas de Calefacción, Ventilación y Aire Acondicionado (HVAC). Se demuestra que, mediante análisis de datos, es posible mantener una detección de fallos eficaz usando menos sensores. De este modo, los costes se reducen significativamente sin sacrificar la fiabilidad del sistema.

Palabras clave: HVAC, sensores, AFDD, optimización de sensores, optimización de costes.

1.Introducción

En los últimos años, el uso de datos ha aumentado exponencialmente, haciendo que las estrategias basadas en estos hayan transformado la gestión y optimización de los sistemas en diferentes areas siendo la automatización de edificios un claro ejemplo. Los edificios más modernos incorporan diferentes sensores (humedad, temperatura, control de compuertas, ocupación, etc.) en sus sistemas de climatización (HVAC). Mediante el uso de estos sensores se puede controlar el confort, uso energético y detección de fallos en el sistema. Sin embargo, la instalación, mantenimiento y almacenamiento de datos de estos sensores puede llegar a ser muy costoso para los usuarios. Este proyecto investiga como reducir el número de sensores utilizados en un sistema HVAC manteniendo una alta precisión en la detección de anomalías. detección de anomalías.

2. Definición del Proyecto

El objetivo principal es analizar el estudio entre la disponibilidad de sensores y la eficacia de la detección de fallos en sistemas HVAC. En concreto, se compara la viabilidad de disminuir el número de sensores manteniendo alta la precisión de los algoritmos de detección. Para ello se realiza un estudio detallado de las métricas de los sensores y un análisis del impacto de eliminar sensores sobre el rendimiento a la hora de identificar fallos

en el sistema. Mediante la realización de este proyecto se da respuesta a las siguientes preguntas:

- ¿Cuál es el número mínimo de sensores que permite una alta precisión absoluta en la detección de fallos?
- ¿Qué técnicas de machine learning minimizan la pérdida de precisión en la detección de fallos con un número reducido de sensores?
- ¿Cuál es la relación entre el número de sensores y la precisión de los métodos de detección de fallos? ¿Cómo se puede caracterizar esto para fundamentar la toma de decisiones sobre inversiones en sensores para HVAC?

Tras responder entre otras a estas preguntas, este proyecto ofrece resultados y soluciones prácticas para los profesionales que podría resultar útil para el desarrollo de sistemas más económicos y eficientes.

3. Descripción del Modelo/Sistema

La metodología usada sigue un enfoque basado en datos, combinando medidas reales de un edificio pequeño en Iowa con simulaciones sobre este. Este trabajo incluye:

- Preprocesamiento y Visualización de datos: depuración de los datos y visualización de ellos para garantizar la calidad y consistencia de estos.
- Análisis de Componentes Principales (PCA): reducción de dimensionalidad para identificar las variables más informativas y observar posibles redundancias.
- Selección de variables: Se utiliza la explicación aditiva de Shapley (SHAP) para identificar cuáles son los sensores con mayor contribución a la detección de fallos en el sistema.
- Algoritmos de machine learning: uso de modelos supervisados (Árboles de Decisión, Random Forest y XGBoost) para evaluar el rendimiento de la detección de fallos bajo distintas configuraciones de sensores.
- Análisis de costes: estimación de ahorro potencial de instalación al emplear un subconjunto optimizado de sensores.

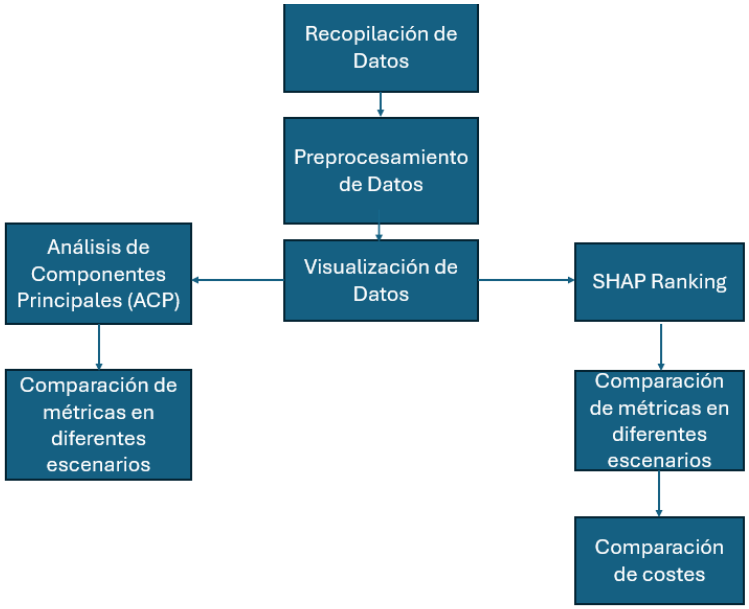

Ilustración 1. Esquema del modelo

4. Resultados

El estudio revela que es posible reducir de forma notable el número de sensores manteniendo alta precisión en la detección de fallos y reduciendo costes. Entre los hallazgos más relevantes se encuentran:

- Identificación de redundancias: PCA y SHAP muestran sensores con información solapada, lo que permite la reducción el número de sensores.
- Rendimiento de detección de fallos: tras utilizar diferentes técnicas de aprendizaje automático supervisado y analizar las métricas sobre un conjunto reducido de sensores, se observa que estas son comparables a las de un conjunto completo. Se observa cómo, con un subconjunto de los cuatro sensores de mejor rendimiento, se captura entre el 98 % y el 99 % de la varianza total de la información (PCA) y una precisión del 99 % en la detección de fallos del sistema.
- Optimización de costes: la reducción de sensores conlleva un ahorro significativo. Es observable como con únicamente los cuatro sensores, el coste total de instalación se reduce un 62.61%

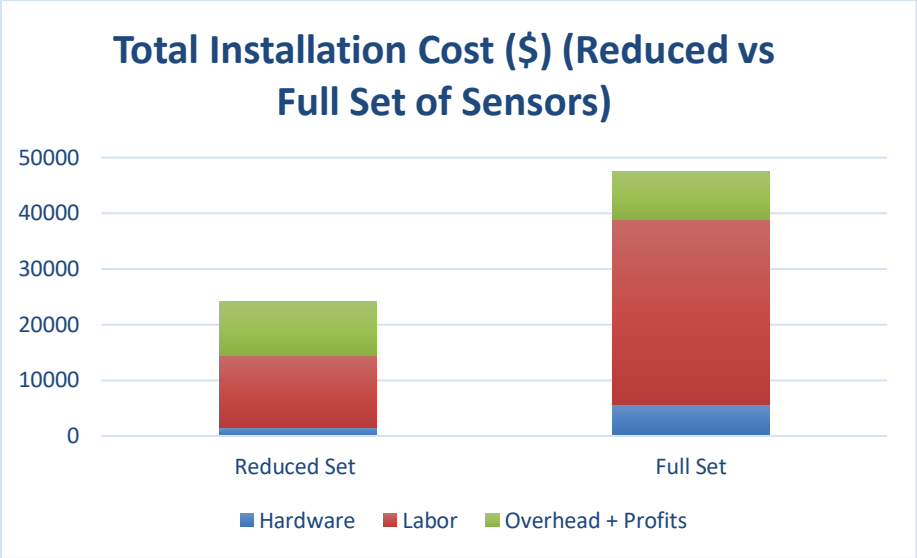

Ilustración 2. Coste Total de Instalación ($) (Número reducido vs total de sensores)

5. Conclusión

Este proyecto proporciona evidencia empírica de que es posible optimizar el número de sensores en HVAC sin comprometer la eficiencia de la Detección y Diagnóstico Automático de Fallos (AFDD). El uso de machine learning y técnicas de análisis de datos permite mantener un sistema fiable a la vez que se disminuyen los gastos. Los resultados obtenidos impulsarán a responsables de edificios y diseñadores de sistemas a replantear las estrategias tradicionales de implementación de sensores, favoreciendo un enfoque más sostenible y económico. El trabajo futuro podría aplicarse a diferentes tipos de sensores, así como a otros campos como puede ser las ciudades inteligentes.

6. Bibliografía

1 **1.** J. Granderson, G. Lin, A. Harding, P. Im, and Y. Chen, "Dataset for building fault detection and diagnostics algorithm creation and performance testing, Version 3," *figshare*, 2020. [Online]. Available: https://doi.org/10.6084/m9.figshare.11752740.v3

2 Z. Chen *et al.*, "A review of data-driven fault detection and diagnostics for building HVAC systems," *Applied Energy*, vol. 339, p. 121030, Jun. 2023. [Online]. Available: https://doi.org/10.1016/j.apenergy.2023.121030

3 K. Trenbath, R. Meyer, K. Woldekidan, K. Maisha, and M. Harris, *Commercial Building Sensors and Controls Systems – Barriers, Drivers, and Costs*, National Renewable Energy Laboratory, Tech. Rep. NREL/TP-6A50-82117, Aug. 2022. [Online]. Available: https://www.nrel.gov/docs/fy22osti/82117.pdf

CHARACTERIZING TRADEOFFS IN SENSOR AVAILABILITY AND FAULT DETECTION EFFICACY

Author: Sanz Barro, Ignacio
Supervisor: Fierro, Gabe
Collaborating Entity: Colorado School of Mines

PROJECT RESUME

In this project the tradeoffs between sensor availability and the fault detection effectiveness in Heating, Ventilation, and Air Conditioning (HVAC) systems are studied. It is shown that it is possible to maintain effective fault detection with data analysis, while using less sensors. Therefore, costs can significantly reduce while maintaining a reliable system.

Key Words: HVAC, sensors, AFDD, sensor optimization, cost optimization

1.Introduction

Over the past years, data driven strategies have changed the perspective of system management and optimization in different fields, where building automation is a clear example. Newer buildings have many sensors such as temperature, humidity, fan status, damper control, occupancy, etc. These sensors are very common in Heating, Ventilation, Air Conditioning (HVAC) systems. These sensors allow for real-time monitoring and control in buildings, thus refining comfort, energy usage and fault detection. However, the cost associated with these sensors are known for being high, not only for its installation cost, but also for the maintenance costs, data storage costs among others. This project proposes an investigation into reducing the number of sensors that can be deployed in a HVAC system, while maintaining a high accuracy in capturing errors at the system.

2. Project Definition

The main goal of this project is to study the tradeoffs between sensor availability and the fault detection effectiveness in HVAC systems. More specifically, the project aims to compare the feasibility of reducing the total number of sensors used in HVAC systems while maintaining a high accuracy of fault detection algorithms. This involves a complete analysis of sensor data and an examination of the reduction of sensors and its effects on overall performance.

The research covers the following important questions:

- What is the minimum number of sensors that can maintain a high absolute fault detection accuracy?
- What machine learning techniques minimize loss of fault detection accuracy with reduced numbers of sensors?
- What is the tradeoff between the number of sensors and the accuracy of fault detection methods? How can this be characterized to inform decision making about investments in sensors for HVAC systems?

By answering these questions among other questions, this study provides actionable insights for professionals that could be helpful for the development of more efficient and cost-effective systems.

3. Description of the Model/System Tool

The methodology used in this project follows a data-driven approach, including both experimental and simulated data from a real small building in Iowa. The work includes:

- Data Preprocessing and Visualization: Initial exploration and cleaning of sensor data to ensure quality and consistency.
- Principal Component Analysis (PCA): Dimensionality reduction to identify the most informative variables and assess sensor redundancy.
- Feature Selection: SHAP (Shapley Additive explanation) analysis is used to identify which features are the ones that contribute the most to identifying faults in the system.
- Machine Learning Algorithms: Implementation of supervision models (Decision Trees, Random Forest and XGBoost) to evaluate fault detection performance under various sensor configurations.
- Cost Analysis: an analysis of potential cost saving is performed if a reduced number of sensors are used.

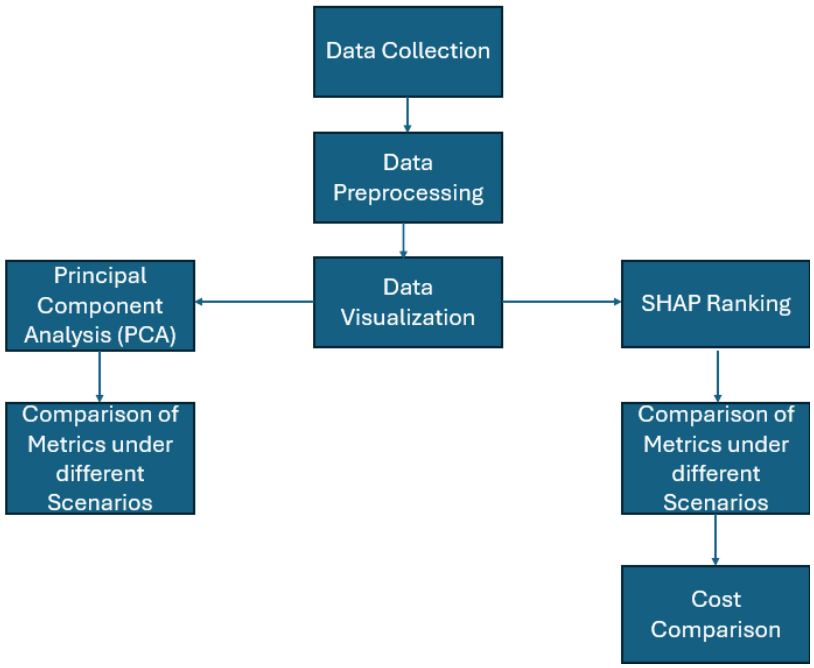

Illustration 1. Scheme of the model

4. Results

After a complete analysis it can be reflected how there could be a significant reduction in the number of sensors while maintaining a high accuracy in detecting faults and reducing costs. Some key findings are the following:

- Redundancy Identification: PCA and SHAP show how some sensors give overlapping information, therefore allowing sensor reduction.
- Fault Detection Performance: After having used different supervised machine learning techniques and having analyzed the metrics with a reduced set of sensors, it can be observed how the metrics are comparable to a full set of sensors. It is observed how with a subset of top four performance sensors is captured 98-99 % of the total information variance and 99% of accuracy on detection faults in the system.
- Cost Optimization: After reducing the sensor set it can be observed how the total installation cost is reduced by 62.61%.

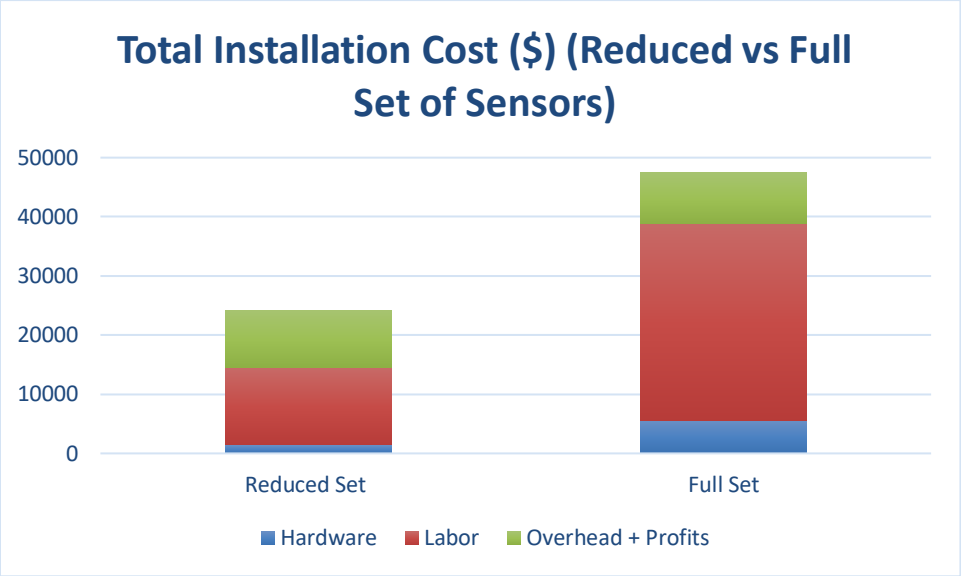

Illustration 2. Total Installation Cost ($) (Reduced vs Full Set of Sensors)

5. Conclusion

This project provides empirical evidence that HVAC systems' sensor number can be optimized while maintaining the efficiency of Automated Fault Detection and Diagnosis (AFDD). It is possible to achieve effective fault detection by utilizing machine learning and data analytics, while decreasing expenses. The findings will make to rethink building managers and systems designers about traditional sensor deployment strategies and point towards a more sustainable and economical way of building management. Future work may be used in different sensor and building types, or in other sectors such as smart cities.

6. Bibliography

1 **1.** J. Granderson, G. Lin, A. Harding, P. Im, and Y. Chen, "Dataset for building fault detection and diagnostics algorithm creation and performance testing, Version 3," *figshare*, 2020. [Online]. Available: https://doi.org/10.6084/m9.figshare.11752740.v3

2 Z. Chen *et al.*, "A review of data-driven fault detection and diagnostics for building HVAC systems," *Applied Energy*, vol. 339, p. 121030, Jun. 2023. [Online]. Available: https://doi.org/10.1016/j.apenergy.2023.121030

3 K. Trenbath, R. Meyer, K. Woldekidan, K. Maisha, and M. Harris, *Commercial Building Sensors and Controls Systems – Barriers, Drivers, and Costs*, National Renewable Energy Laboratory, Tech. Rep. NREL/TP-6A50-82117, Aug. 2022. [Online]. Available: https://www.nrel.gov/docs/fy22osti/82117.pdf

Memory Index

Index of Figures

Index of Tables

Chapter 1. INTRODUCTION

Over the last decades and due to technology, data has become an inseparable part of our daily lives. Data use has kept growing over the last years, shaping the way it is understood, managed and optimized systems across different industries. As can be seen in the graph provided below, data creation has experienced exponential growth over the last ten years.

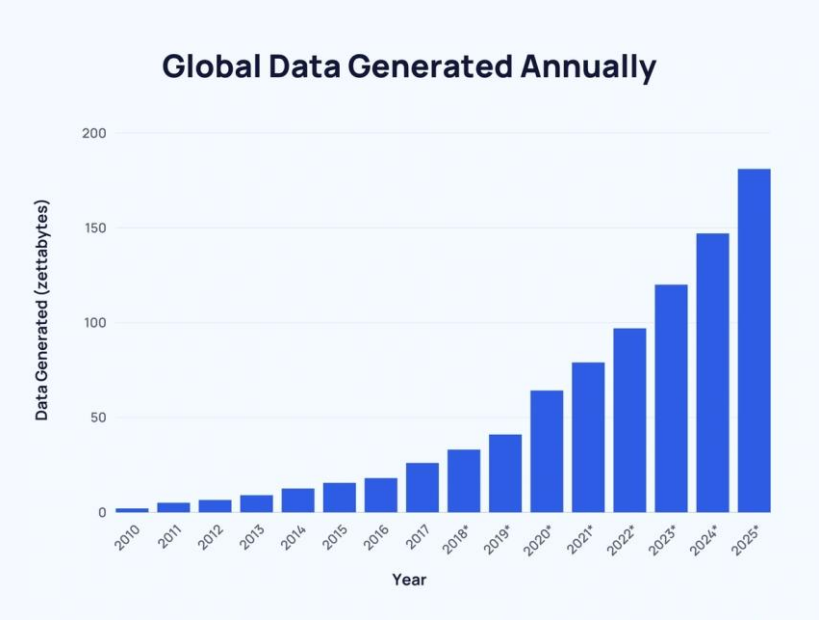

Figure 1. Global data generated annually in the last 15 years. Source: [15]

Data-driven strategies have enabled industries such as Financial, Healthcare or Energy to change their traditional methods of coming up with decisions. Data-driven strategies are enabling companies to make more accurate decisions based on data, reducing inefficacies and developing predictive analysis that were unimageable decades ago.

Moreover, smart cities and smart buildings have become more popular during the last decades. Nowadays, buildings are between other items, composed of different sensors.

Sensors such as humidity sensors, air quality sensors, occupancy sensors, pressure sensors or temperature sensors are used in multiple buildings systems.

Furthermore, there has been an increased awareness about energy efficiency and the usage of these sensors help to optimize building systems and to reduce unnecessary energy use.

The integration of sensors into building systems, mainly in Heating, Ventilation, and Air Conditioning (HVAC) systems, has been a game changer in the operation and efficiency. These sensors provide real-time data, which helps buildings to adapt the way they perform depending on the environmental conditions, user needs or other factors. This data-driven adaptability is crucial in modern building sets, where comfort, energy efficiency and fault detection are fundamental.

However, there are several inconveniences that should be considered to provide a better system. It is proved that these sensors are associated with a high cost. The cost of these sensors does not only depend on the installation cost, but it also depends on the maintenance cost, data storage cost and unexpected costs. Additionally, handling larger amounts of data can significantly increase the complexity of computation.

With this project, the goal is to analyze the functionality of some of the sensors included in the dataset and evaluate whether it is feasible to reduce the number of sensors without compromising fault detection accuracy or overall system performance. This potential reduction aims to maintain a high diagnostic effectiveness in detecting faults while optimizing costs.

Furthermore, the results and conclusions of this project could allow experts in the industry to take a deeper look at their system and might redefine how their systems work. It could be supervised the way in which data is used in different areas such as smart cities, as the amount of data used could be surplus and the cost associated with it higher than what it should be.

In the last few decades, Automated Fault Detection and Diagnosis (AFDD) has been recognized as a critical component of building management system optimization. AFDD technology uses sensor readings to identify early system failure or inefficiencies, preventing

future faults and improving overall performance. Though the effectiveness of such technologies is supposed to be inversely proportional to the amount of data gathered, this has brought with it a widespread misconception that placing more sensors necessarily translates to increased fault detection. While intuitive as a concept, such an argument provokes substantial questions regarding whether large numbers of sensors are necessary at all, especially given the practical and financial aspects. This project attempts to challenge that generalization through research on the relationship between sensor availability and fault detection performance.

The objective is to use machine learning and data analytics across an operational scenario continuum to determine if the same, or very similar, levels of performance can be attained with fewer sensors. By exploring these trade-offs, it is hoped to facilitate the technical evolution of AFDD systems and support more affordable and sustainable alternatives for smart buildings and infrastructure.

For the realization of this project, real and published data from the Lawrence Berkeley National Laboratory (LBNL) has been used. This laboratory, which is also known as Berkeley Lab, is a United States (U.S.) Department of Energy (DOE) national laboratory managed by the University of California. This institution is recognized worldwide for creating advanced tools for scientific discovery and work-on problems of great scales, especially in areas such as energy, healthcare and the environment.

In particular, the data set used in this study originates from HVAC systems which are being used in real-world conditions, where one of their variables indicates if there is an error or not in the system. This variable provides valuable insights, which with the use of programming tools could identify where the error is at. Furthermore, it will allow the evaluation of how accurate faults can be detected based on sensor readings.

Beyond the technical aspects of the project, the sensors' cost is also studied and how a reduction in those could reduce the total installation cost.

The planning for the development of this project is divided into different intercorrelated stages:

- Literature revision: An extensive review of academic literature was done to gain a relevant and in depth understanding of how the HVAC systems work and what the existing knowledge of the area was. Moreover, I conducted a depth analysis of the different data sets available. Research was mainly conducted through reading existing studies in the sector, many of them being from Lawrence Berkeley National Laboratory (LBNL) [12].
- Data collection and Preprocessing: researched various datasets and finally chose one (with experimental and simulated data) which could detect when there are faults in the HVAC system. Preprocessing, which is fundamental in all data analytics programs, has been done.
- Model Development and Testing: Principal Component Analysis (PCA) has been realized to gain a better understanding of which variables give the most information for the variance of the dataset. Moreover, with different Machine Learning techniques such as Random Forest, Decision Trees and XGBoost, it has been studied which variables had a better performance for different metrics (accuracy, precision, recall, F1 Score). These supervised techniques are seeking to understand which variables target better our target variable.
- Analysis and Visualization: with different combinations of sensors, it has been analyzed the tradeoffs between sensor availability and detection performance.
- Economic Impact: it has been studied which would be the economic impact of reducing different sensors while maintaining a high accuracy in the model. More papers and blogs have been read during this period, as the information differs between different sites.
- Recommendations and Framework Development: findings have been put together and suggested conclusions and future work on this area. Moreover, some of the findings could be used in other sectors, where the use of data is also abundant.

To organize and visualize the progress of the project, a diagram of Gantt has been elaborated which shows clearly and concisely the different stages and activities of the project. Diagram of Gantt is a world-wide tool used in the planification of projects, which allows us to establish a detailed schedule and to manage efficiently the resources and time.

In the following figure, it is presented the diagram of Gantt of this project, where detailed presented the planned tasks, time schedule and the initiation and finalization of the corresponding tasks. This diagram provides an overview of the project schedule, making it easier to identify potential overlaps, delays, or scheduling issues, allowing timely corrective action to be taken.

Tasks	Before February	February	March	April	May
Get informed about HVAC and how to tackle the problem					
Acquire knowledge about how do the sensors work and AFDD					
Data collection and preprocessing					
Use different ML algorithms to train and test the database					
Start modifying the number of sensors to observe the change in accuracy					
Tradeoffs between sensor availability and detection performance, provide visualizations of the different metrics					
Final Paper and Presentation					

Table 1. Diagram of Gantt

Chapter 2. DESCRIPTION OF THE TECHNOLOGIES

Heating, Ventilation, and Air Conditioning (HVAC) systems are critical infrastructures in both residential and commercial modern buildings. The main aim of these systems is to keep a comfortable indoor temperature, reduce energy waste and keep different measurements as expected.

Moreover, there are plenty of HVAC systems that are used in different zones. Some of these types include rooftop units, single-duct air handler units, dual-duct air handler units, variable air volume boxes, chiller plant, boiler plant or fan coil units. In each commercial or residential building, if they are used, one system or various types of systems could be used.

To ensure the reader's correct understanding of the work, the different main components that make up HVAC systems, as well as their use, will be provided.

2.1 HVAC MAIN COMPONENTS

To understand how the HVAC systems work, each main component must be understood independently and what each component contributes to the HVAC system. HVAC systems vary in complexity, but the main components are the following.

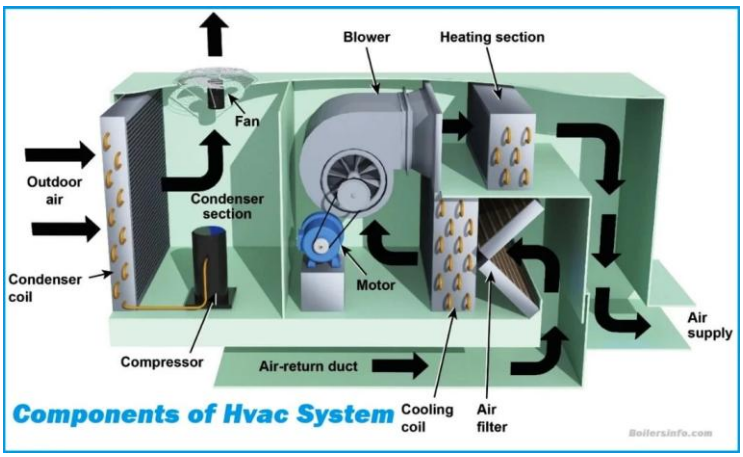

Figure 2. Components of HVAC System [13].

2.1.1 AIR TEMPERATURE SENSORS

These sensors are aimed to check a specific temperature in a concrete zone. These sensors include supply, return, outdoor and mixed temperature. Supply sensors measure the temperature of the air which is delivered to the air by the Air Handling Unit. Return sensors monitor the temperature of the air returning from the occupied space. Mixed Air Temperature captures the mixed temperature from the return air and the outdoor before conditioning. Outdoor Air Temperature Sensors capture the temperature outside of the building, being a key part to deciding whether the building should activate a heating, cooling or none of those status.

2.1.2 DAMPERS AND VALVE

- Return Air Damper Control Signal: It is in charge to control how much return air could go back into the system, being their values in a range of 0(fully closed) and 1(fully open)
- Cooling Coil Valve Control Signal: Controls how the valve would open so that cool water could flow, being their values in a range from 0 (no cool water flow) to 1 (fully open cool water flow).

- Heating Coil Valve Control Signal: Controls how the valve would open so that hot water could flow, being their values in a range from 0 (no hot water flow) to 1 (fully open hot water flow).

2.1.3 AIRFLOW AND PRESSURE MEASUREMENT

- Supply Air Duct Static Pressure: measures the pressure through the supply duct, so that it can ensure an efficient air distribution.

2.1.4 FANS

The air supply fan overseas pushes air to the building's ductwork, making sure air reaches every occupied zone. The return air fan oversees drawing air back from space, keeping air pressure and pressure balance

2.1.5 DUCTWORK

Are the spaces from where the air is moved from one place in the building to another. It is a fundamental part of these systems as without ductwork the air could not move from one place to another.

Moreover, there are different key components in these systems such as the occupancy mode, which will lead the systems to work differently depending on the day and time of the day in which the unit is running on.

2.2 FAULT DETECTION ALGORITHMS IN HVAC

Fault detections in HVAC systems are critical to maintain a building operation as it should. The goal of fault detection in HVAC systems is to identify anomalies in buildings to fix the anomalies as fast as possible to keep the operation of the buildings as they should. These faults include dampers stuck, compressor faults, fan anomalies or thermostatic failures

among others. Detecting such anomalies early is fundamental for maintaining system reliability, energy savings or effective maintenance.

Different techniques are used to detect these faults. These techniques vary depending on the level of knowledge integrated into the algorithm. The different categories are the following:

- White-box models: have complete knowledge about the engineer principles the system is using and what is the HVAC system behavior. These models lean on different formulas which explain the behavior of the system. It considers thermodynamic equations, heat transfer principles or fluid dynamics to detect anomalies. Although they usually provide high accuracy for detecting faults, these formulas are different for each specific building and the HVAC system used and is very time consuming to develop.
- Grey-box models: it is a mix of both white-box and black-box models. In this model, several components or formulas are grouped into a larger subset. With this subset, which requires less knowledge than in the white-box model, data analysis is used to determine whether there are errors in the system. This model is more flexible than the white-box model but is not typically as precise when calculating errors or interpretability.
- Black-box model: this model is fully data-driven and does not require any prior knowledge of how the system operates. These models rely on historical data and with the use of machine learning techniques aim to find faults. These models are very scalable and effective in large or complex systems but lack interpretability and do need a large data set to be trained.

During this project black-box models were used to identify patterns, Patterns through machine learning techniques learn from the data set to find out where the error comes from.

These Machines learning techniques could be broadly categorized into supervised or unsupervised learning methods, depending on the input data used and the desired output.

- Supervised methods: make use of labeled data to train algorithms. From the data, it is known whether it is an input feature and what is the corresponding output label (fault or not fault). The model learns the relationship between inputs and known outputs, so that this knowledge can be used in unseen, new data. Decision Trees, Random Forest or XGBoost algorithms are popular supervised models which are going to be used in this project.
- Unsupervised methods: do not use labeled data. Their aim is to find hidden patterns or structures in the data. These methods are less effective when the output is already defined. Common unsupervised algorithms include K-Means or Clustering.

In this project, as mentioned above supervised methods were selected. These methods are used as a dataset containing an explicit label that allows the user to know when there is a fault in the system. This means that the models could be trained with direct feedback as to how accurate the prediction was, resulting in models that were more accurate and interpretable related to guided fault condition identification.

Chapter 3. STATE OF THE ART

The building field has been going through a significant transformation due to technological advancement in the last decades. The advancement in sensors usage and the increased number of sensors have generated an out of the ordinary volumes of data from buildings systems like HVAC. This data with the use of Automated Fault Detection and Diagnosis (AFDD) is very valuable to identify and fix operational anomalies in these complex systems. These techniques play a fundamental role in getting buildings that are energy efficient, reduce costs, maintain buildings users comfortable and reduce the lifespan of the HVAC equipment. [1].

In this context, HVAC systems can help to reduce significantly energy consumption all over the world, as they are used in areas where energy consumption is very high. According to different studies, it can be observed how approximately 15-30% of energy may be wasted due to building systems faults and improper controls [2]. Therefore, an outstanding performance of AFDD methods in HVAC systems could allow decreases the total amount of energy consumption.

There are numerous ways in which faults could be manifested in HVAC systems. Minor errors such as subtle sensor drifts and calibrations could lead to a fault, or major errors such as stuck dumper or a malfunctional in a fan could also lead to errors. These errors could lead to worse performance, affecting energy consumption. Therefore, a rapid and accurate detection of these errors could mitigate the negative consequences associated with these malfunctions. [1]

Historically and due to the lack of data techniques, AFDD approaches were mainly based on knowledge of the systems and physical models. White-box models, for example, use intricate concepts from fluid dynamics and thermodynamics to model the optimal behavior of a system and spot deviations from what is expected to identify faults. Despite the high accuracy of these models, they are usually very complicated, needing a high level of

technical work and detailed system specifications. Taking all of it into account, it is hard to scale and modify to account for the inherent diversity of HVAC systems across different buildings [3]. Rule-based models (grey-box), which use slightly less knowledge, but are still a knowledge-driven approach, use a set of predetermined rules based on expert knowledge. This approach indicates when a sensor reading is out of the expected range of values or when it does not meet with a specific logical condition. Although these models are easier to implement than physical- models (white-box models), they may have trouble identifying new or complex fault patterns, and their efficacy is restricted by the quality and completeness of the rule base.

The exponential increase in the number of sensors and therefore the multiplication of the data has started a shift towards data-driven AFDD methods. With the use of machine learning techniques and statistics these methods learn patterns from normal and malfunctioning operations from historical data. These purely data-driven models (black-box models) have increased exponentially their popularity as they are able without any deep understanding of the system and its physical principles to deal with complex and non-linear relationships within the data. [3,4]

When in a labeled dataset it is available whether there is or not an error in the system, supervised learning techniques are frequently used in data-driven AFDD. In order to create classification models that can predict existence and sometimes the type of failure based on real-time sensor readings, methods such as Decision Trees, Random Forests, Support Vector Machines (SVMs) and Gradient Boosting Machines (such as XGBoost) are trained on this labeled data. These methods have shown a great deal of effectiveness in locating recognized fault patterns that have been discovered through historical occurrences. [4].

On the other hand, unsupervised learning methods are used when there is no labeled fault data. Unusual data points or patterns that substantially depart from typical operating conditions can be found using clustering algorithms and anomaly detection techniques, which may hint to previously unknown faults [4].

Although data-driven AFDD methods are great dealing with great volume of data, it is very fast and can integrate different sensors, these methods do also present significant challenges. Data quality issues, such as missing values, redundant measurements, noise or sensor failure lead to loss of valuable information, impacting the performance of AFDD algorithms. Moreover, the high dimensionality of datasets with many sensor inputs can increase significantly the model complexity and computational cost and lead to overfitting in machine learning models [5].

After having taken all this into account, there has been an increased study to optimize the use of sensor data for AFDD, aiming to be the topic of reduction and selection of the sensors. Although, more data is associated with better fault detection, this assumption is being challenged by the practical realities of cost and data management complexity. Moreover, the implication of using more data or sensors is linked with a higher cost. This cost does not only include the initial installation and hardware cost, but also includes expenses related to processing infrastructure, maintaining costs or data storage among other costs [6].

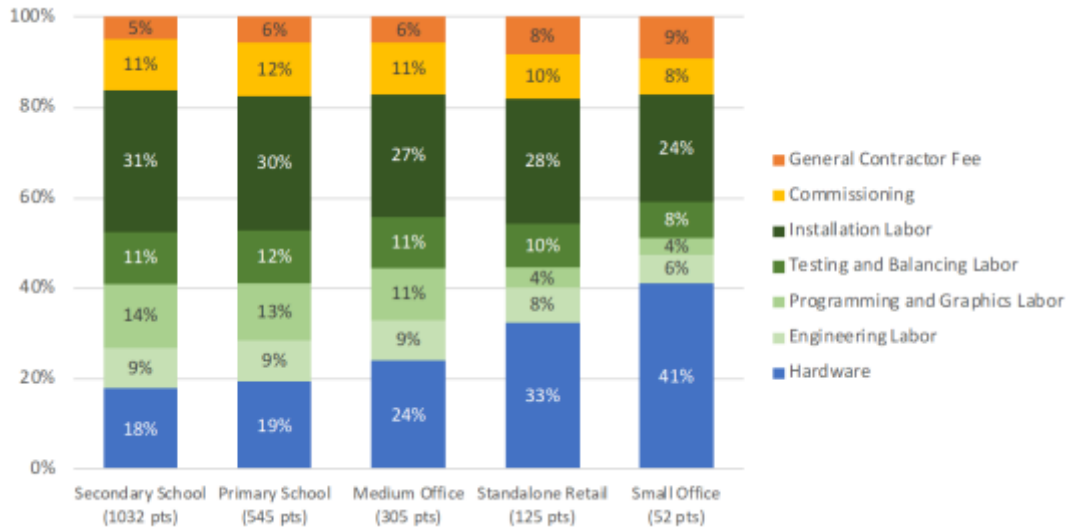

Figure 3. Cost breakdown comparison for complete control systems Source: [6].

Therefore, and to reduce costs and redundancy, identifying the most informative group of sensors is an important step towards developing cost-effective and efficient AFDD systems. There are different techniques that can be used for this purpose. Principal Component Analysis (PCA) is a great dimensionality reduction technique, which can be used to transform the original data det into a lower dimensionality set of principal components while maintaining most of the data's variance [7]. The aim of the Principal Component Analysis is to provide useful insights into which combinations of sensors contribute to the overall system variability, and this information may be useful for identifying fault detection. Moreover, there are other techniques, named feature selection techniques, that aim to identify and select which mixture of the original sensors are the most appropriate for predicting fault detection. These feature selection techniques are divided into methods such as filter-based, wrapper-based and embedded-based techniques. The filter-based method consists of selecting features based on statistical measures without involving any machine learning model. The wrapper-based method trains and tests different subsets of data to find the best combination. Lastly, the embedded-based method performs feature selection as part of the model training process [8].

It is essential to conduct research on the tradeoff between the AFDD performance and the number of sensors. There are plenty of investigators and studies that have investigated the effects of shrinking the sensor set on important performance metrics such as accuracy, precision, recall and F1 score for fault detection. It is true that a reduction in sensors without any knowledge of that sensor reduction can potentially lead to a degraded performance, but an intelligent sensor selection based on different data analysis approaches mentioned above may be able to maintain or even approach the performance of systems with a full sensor suite [8,9]. The economic advantages of such optimization are clear: AFDD solutions can become more accessible and affordable as they will lead to lower sensor costs, less installation costs and less data management overhead. This reduction of sensors will speed up in small and medium size buildings as they are the ones with higher cost sensitivity [6].

Although there has been significant progress in the field, there are still several research gaps. Firstly, despite it is known that with a reduced subset of sensors, there are lower costs, a detailed and quantification of the economic impacts linked to the impact on various AFDD performance metrics needs further investigation. There is still a need for specific research using actual data to illustrate this trade-off. Additionally, research is still ongoing to determine the stability and transferability of reduced-sensor AFDD models and sensor selection procedures across various HVAC system types, manufacturers, and building operational profiles [5]. The implementation of standardized methodologies for comparative AFDD performance with different sensor subsets is also fundamental for comparing different methods objectively.

Real world HVAC systems are fundamental to further investigation in data-driven AFDD and sensor optimization. In this sense, the Lawrence Berkeley National Laboratory (LBNL) is a pioneer. The LBNL institution provides different public datasets from different buildings which include operational data from different HVAC systems of both normal and faulty detection [10]. These datasets, which presents frequently ground truth labels for errors, are essential tools for AFDD algorithm and sensor selection training, validation and benchmarking. These high-quality real-world data are the key for developing workable and efficient AFDD solutions.

To sum up this section, the state of art in HVAC is AFDD is characterized by a shift towards data-driven approaches with the help of machine learning techniques as there is an increasement in the availability of sensor data. The main benefits of this shift are moving towards a more energy efficient world and reducing costs associated with the reduction of sensors, there are different challenges related to the high volume of data and which sensors are the most appropriate to reduce to gain sensor optimization. Several research points out useful techniques such as PCA or feature selection to find which sensors are the most informative and to analyze the trade-off between sensor count and fault detection performance.

However, a comprehensive economic analysis of sensor reduction tactics and an exploration of these methods in different HVAC systems are essential to promote the broader implementation of cost-effective and precise AFDD solutions. With this project, after analyzing real-world data to quantify the impact of sensor reduction on fault detection accuracy, it is studied which would be the economic impact of this reduction.

Chapter 4. PROJECT DEFINITION

4.1 JUSTIFICATION

Over the past years, sustainability, energy efficiency and being operational reliable have become fundamental goals to develop and manage modern buildings. Among different building systems, HVAC (Heating, Ventilation and Air Conditioning), are crucial to obtaining the goals mentioned above, as this system is one of the largest energy consumers in commercial buildings. Moreover, failures in HVAC components affect directly the indoor conformity of the people and items inside the building, increase the costs of the building's maintenance, increase energy consumption and the components deteriorate and therefore reduce their lifespan.

The increasing popularity of smart buildings and sensors has resulted in significant growth of data-driven methods to improve HVAC systems' performance. These modern technologies make us able to collect great amounts of data from sensors. This data is used to acquire information from the sensors and with the data analyzes the system's behavior in real time. However, sometimes it is still difficult to transform this raw data into useful insights, though, particularly when dealing with intricate systems and a limited technical system.

Traditional fault detection methods typically involve white-box and grey-box models. As mentioned in the previous chapter, the white-box model (physics-based model), requires a deeper understanding of the system under study and is time consuming to build. Moreover, the grey-box model (rule-based system), although it simplifies the physic-based model it often does not scale as much or is unable to detect combined faults. In addition to this, nowadays, the systems provide many sensors, which increases both the cost and the complexity of the solution under these two models. With this casuistry, there is a need to create a more flexible, scalable and efficient model, which can be applied to different buildings and HVAC systems with just minor changes.

There is a misconception in the use of sensors in HVAC systems and other systems. Many people associate the use of more sensors with better diagnosis and monitoring. While this conception is usually true, it should be considered other issues associated with an excessive number of sensors. Some of these issues are the high installation and maintenance cost, cost of data storage and might have an increased complexity for data interpretation. This consideration contradicts also the principles of cost-efficiency and sustainability that smart buildings aim to promote.

What brings me to develop this project is to address those challenges by developing a data-driven model. With the usage of supervised machine learning and data where the faults are labeled, this project is both effective and practical to implement. Data classifies the operating conditions and detects faults based on the labeled data. It studies how absolute accuracy changes if it is used a reduced number of sensors. The data sets used are simulated and experimental used in a building environment, developed by the U.S Department of Energy and Drexel University. It provides detailed information about different scenarios, which allows models to be trained with confidence as ground-truth labels are known.

4.2 OBJECTIVES

This project's main objectives are both technical and practical, with the purpose of creating a solution that can be used in the real world. Although a core part of the project is the usage of supervised learning techniques, this project also evaluates the economic implications, exploring dimensionality reduction and testing generalizations.

The objectives of the project are the following:

- Develop a fault detection system using supervised machine learning algorithms (Decision Tree, Random Forest, XGBoost).
- Evaluate the models using simulated and experimental HVAC data with known ground truth for faults.

- Apply Principal Component Analysis (PCA) to identify which variables tell you the most out of the data set.
- Apply SHAP ranking to observe which variables are the ones that are related the most with identifying anomalies (ground-truth).
- Quantify the relationship between the number of variables used and the performance of the models.
- Determine whether a reduction in the number of sensors can still maintain acceptable levels of absolute accuracy and fault detection reliability.
- Assess the economic impact of the reduction of the number of sensors in real-world buildings systems.
- Identify different configurations of sensors that could be used to obtain high accuracy and a reduction in costs.

4.3 BROADER VALUE OF THE PROJECT

With this project, I aim to provide insights into HVAC systems developers for them to rethink how HVAC systems should be managed. By providing that it is possible to achieve reliable fault detection with a reduction of sensors, the project challenges current assumptions and opens the door for more efficient, sustainable and cost-effective designs. The development of the project can be used not only in HVAC systems, but also in different approaches such as lighting systems, security systems or even other sectors not related to buildings where great amounts of data is used such as smart cities.

Additionally, the project is completely in line with global trends including the expansion of smart infrastructure, the requirement for predictive maintenance and environmental objectives related to energy efficiency. A system that minimizes sensor usage, identifies issues early, and is interpretable for non-experts to understand offers a distinct value proposition from a technical and business perspective.

With the mix of academic methods (PCA, SHAP ranking, supervised machine learning techniques, simulation analysis) with real-world business implications (cost planning, scalability, interpretability), the project offers a wide view of how technology can serve as the bridge between innovation and practicality.

Chapter 5. DEVELOPED SYSTEM/MODEL

5.1 INTRODUCTION

To start developing a system to obtain the features which are the most important to predict accuracy, it is fundamental to understand how does the prediction work. The development of the system has followed a similar approach to the described in the figure above, a general data-driven FDD process [5].

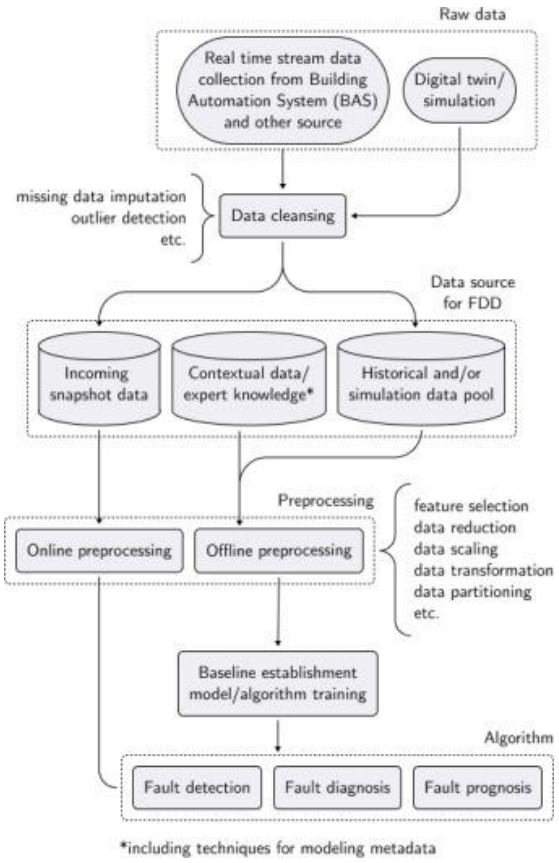

Figure 4: General data-driven FDD process. Source: [5]

The process in which the detection and prediction of faults in buildings systems like HVAC with just data is shown in the diagram above.

The process starts with the collection of data from simulations or real-time sensors and then cleaning missing values. After that, the data is categorized into three different categories: incoming snapshot data, expert knowledge and historical/simulated data.

It then undergoes processing, such as scaling or feature selection, to get it ready for analysis.

Lastly, machine learning models are taught to identify issues, know where failure comes from, and even forecast potential problems in the future.

5.2 GET TO KNOW THE DATA

Depending on which dataset and the building under study, variables vary. In each dataset, it should be understood what each variable is used for, and which is the measurement of each variable.

The data under scope is provided from the US Department of Energy [10]. The data has been collected and simulated by Drexel University for a small commercial building in the state of Iowa, United States.

The building being examined is part of a test facility designed to evaluate different energy-saving strategies under real and controlled conditions. The building's setup allows independent testing from identical zones. The building makes use of three air handling units (AHUs). AHU-1 is used in the common areas. On the other hand, AHU-A and AHU-B serve four zones each, both identical. The data used for experiments and simulations is provided by AHU-A.

Moreover, it is worth mentioning that under normal conditions, the east and west-facing rooms were operated without any imposed faults. On the other hand, when establishing data under faulty condition, the east and west-facing rooms were operated deliberately imposed faults while the south room ran normally [11].

The distribution of the building studied in this project is provided below and is also provided by which Air Handling Unit (AHU) is served each room.

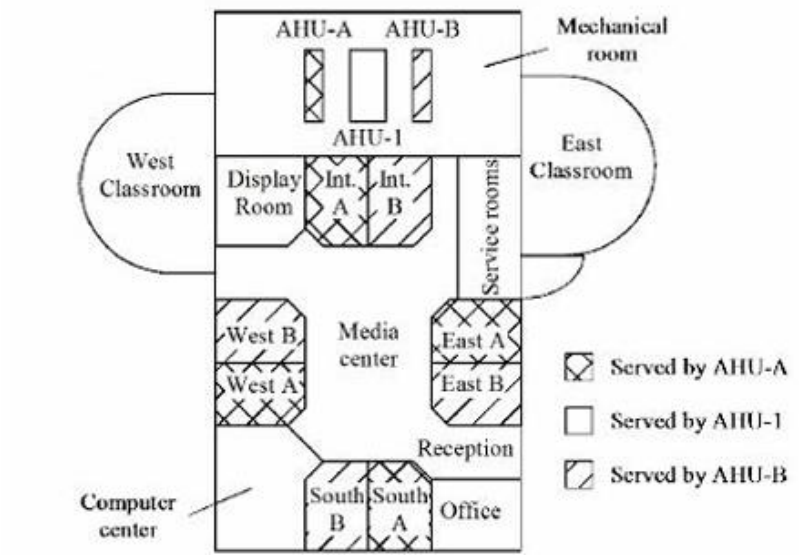

Figure 5. Distribution of the building under study. Source: [10,11]

5.3 *DATA VARIABLES*

In this section, the variables included in the dataset are presented. It is fundamental to understand in depth the variables, what each variable is describing and the units or measurements associated with them. A correct understanding of the variables can provide useful insights to identify anomalies or conclusions, which is essential for any data-driven analysis. Some variables may reflect direct sensor readings (like temperature or humidity), while others represent control signals or calculated values. Understanding this distinction is key when applying different machine learning algorithms or analyzing data.

The data set used has the following variables:

Data Point Name	Description	Unit
AHU: Supply Air Temperature	Measured AHU supply air temperature	°F
AHU: Supply Air Temperature Set Point	AHU supply air temperature set point	°F
AHU: Outdoor Air Temperature	Measured AHU outdoor air temperature	°F
AHU: Mixed Air Temperature	Measured AHU mixed air temperature	°F
AHU: Return Air Temperature	Measured AHU return air temperature	°F
AHU: Supply Air Fan Status	AHU supply air fan status; 0-off, 1-on	--
AHU: Return Air Fan Status	AHU return air fan status; 0-off, 1-on	--
AHU: Supply Air Fan Speed Control Signal	AHU supply air fan speed; ranges from 0 to 1; 0 - fan speed is 0%, 1 - fan speed is 100%	--
AHU: Return Air Fan Speed Control Signal	AHU return air fan speed; ranges from 0 to 1; 0 - fan speed is 0%, 1 - fan speed is 100%	--
AHU: Exhaust Air Damper Control Signal	Control signal for AHU exhaust air damper; ranges from 0 to 1; 0 – damper should be fully closed, 1 – damper should be fully open	--
AHU: Outdoor Air Damper Control Signal	Control signal for AHU outdoor air damper; ranges from 0 to 1; 0 – damper should be fully closed, 1 – damper should be fully open	--
AHU: Return Air Damper Control Signal	Control signal for AHU return air damper; ranges from 0 to 1; 0 – damper should be fully closed, 1 – damper should be fully open	--
AHU: Cooling Coil Valve Control Signal	Control signal for AHU cooling coil valve; ranges from 0 to 1; 0 – valve should be fully closed, 1 – valve should be fully open	--
AHU: Heating Coil Valve Control Signal	Control signal for AHU heating coil valve; ranges from 0 to 1; 0 – valve should be fully closed, 1 – valve should be fully open	--
AHU: Supply Air Duct Static Pressure	Measured AHU supply air duct static pressure	psi
AHU: Supply Air Duct Static Pressure Set Point	AHU supply air duct static pressure setpoint	psi
Occupancy Mode Indicator	Indicator if the system operates in occupied mode; 1-occupied mode, 0-unoccupied mode	--
Fault Detection Ground Truth	Indicator if there is a fault present during the day; 0 – unfaulted, 1 - faulted	--

Table 2: Variables Description [10]

Although the table provided above does not point out which are the differences between the different occupancy modes, the main difference is the time in which the systems operate. The occupancy modes are the following:

- Occupied Mode: Monday-Sunday from 6am-6pm.
- Unoccupied Mode: Monday- Sunday the rest of the time.

As mentioned previously, this data is collected from the US Department of Energy and the units used are not the ones used in the International System of Units. The following formulas convert these units into international units.

$$°C = \frac{5}{9} * (°F - 32)$$

$$Bar = psi * 0.0689474$$

5.4 DATA PREPROCESSING AND VISUALIZATION

One important step when handling data is to enhance a better understanding of data by representing it in a visual way.

Firstly, the data has been downloaded from the Lawrence Berkely National Laboratory [12]. Both the experimental and simulated dataset have been downloaded to observe different scenarios.

Both datasets have nineteen variables (columns), which have been described in Table 1. On the other side, the simulated dataset has more rows than the experimental, the number of rows being 37440 and 21600 respectively. Moreover, both scenarios do not have any missing values.

Then, the target variable “Fault Detection Ground Truth” has been analyzed. This variable lets us know whether there is a fault in the system or if there is not. It has been analyzed in both simulated and experimental datasets.

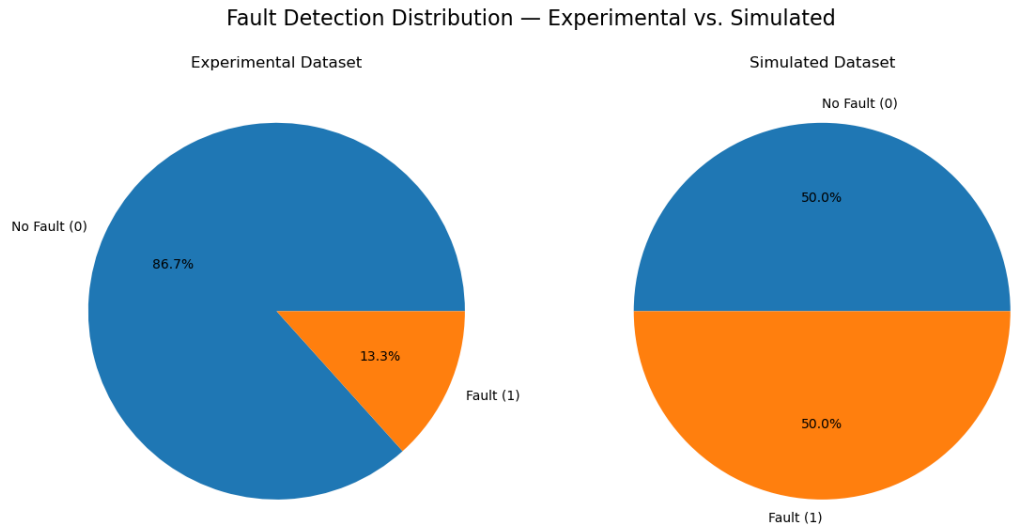

Figure 6.Fault Detection Distribution — Experimental vs. Simulated

It can be observed how the experimental data has less faults than the simulated dataset, where the experimental dataset has only 13.33% of faults, while the simulated dataset has 50% of faults. It is important to point out that the simulation dataset is not what happens in the real world but forces the system to work under different scenarios. Moreover, it is studied whether there is a difference in the fault detection during different occupied modes, and as it is perceived in the graph below, there are no differences.

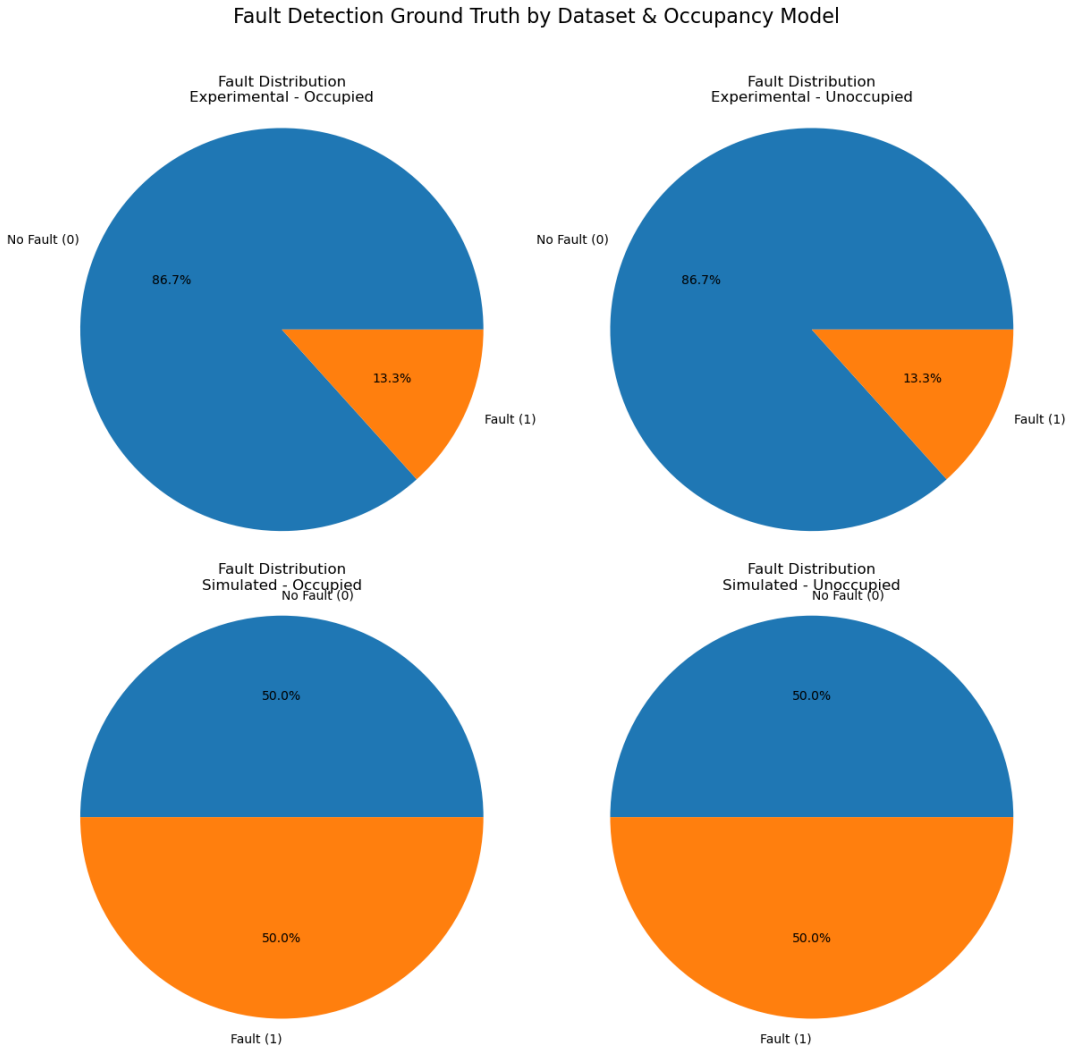

Figure 7.Fault Detection Ground Truth by Dataset & Occupancy Model

From the illustration provided above, it is worth noting that it operates in the same way whether it operates during occupied or unoccupied mode.

Furthermore, it provides different visualizations which represent different variables in the different datasets. To obtain a better understanding of the difference between the two datasets, the data is counted as a percentage of the total rows in each dataset.

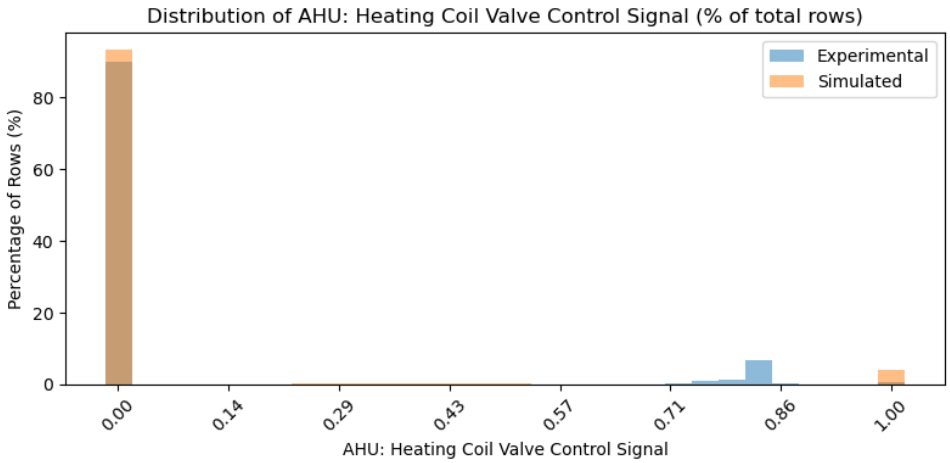

Figure 8. Distribution of AHU: Heating Coil Valve Control Signal'

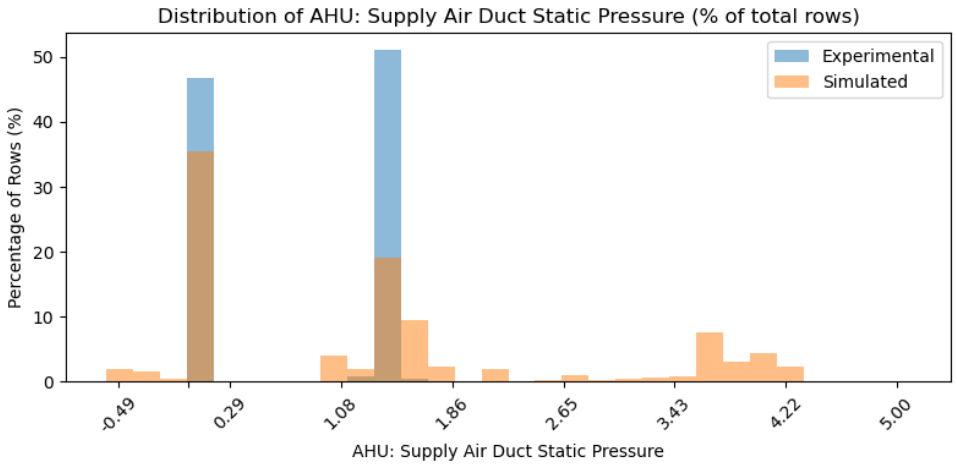

Figure 9: Distribution of AHU: Supply Air Duct Static Pressure

From Figure 8, it can be viewed how the AHU heating coil valve, although approximately 90% of the time is closed, it also modulates between 60% to 90%, but on the other hand the simulated data stays approximately 90% closed, with only occasional full openings. Figure 9 shows how in the experimental data the air duct pressure supply is more static with just two values, whereas the simulated dataset jumps to different values, up to 4.30 psi, producing

unrealistic swings. These differences highlight the simulator's failure to capture dynamics from the real system.

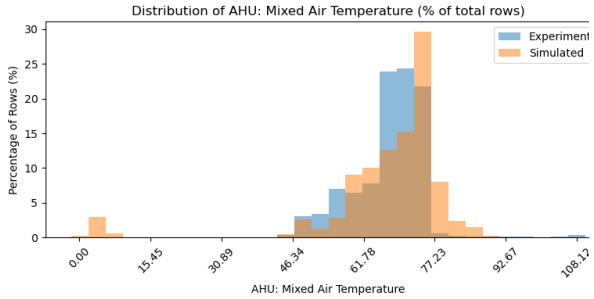

Figure 10: Distribution of AHU: Mixed Air Temperature (ºF)

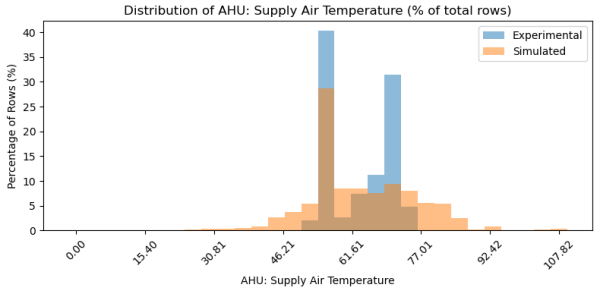

Figure 11: Distribution of AHU: Supply Air Temperature (ºF)

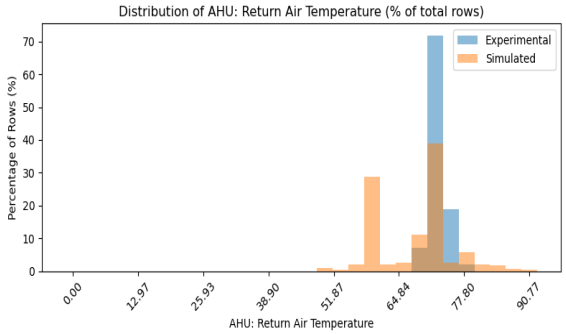

Figure 12. Distribution of AHU: Return Air Temperature (ºF)

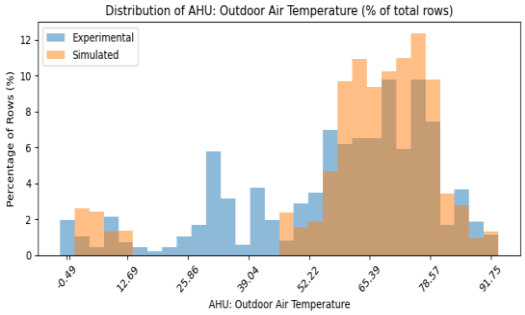

Figure 13. Distribution of AHU: Outdoor Air Temperature (ºF)

From the four different graphs provided above, which represent the Air Handling Unit supply, mixed, outdoor and return temperature, it can be observed how normally there is a wider range of temperatures on the simulated data than on the experimental data, except with

the outdoor temperature. This observation aligns with the idea that simulated data changes reproduce different behaviors of the system. Moreover, the illustrations show how the supply, mixed and outdoor temperatures are lower overall than the return air temperature. This observation makes sense as the return temperature is the range of temperature monitored to be depending on the system operating mode. This temperature is the one that is perceived by the users that are inside the building.

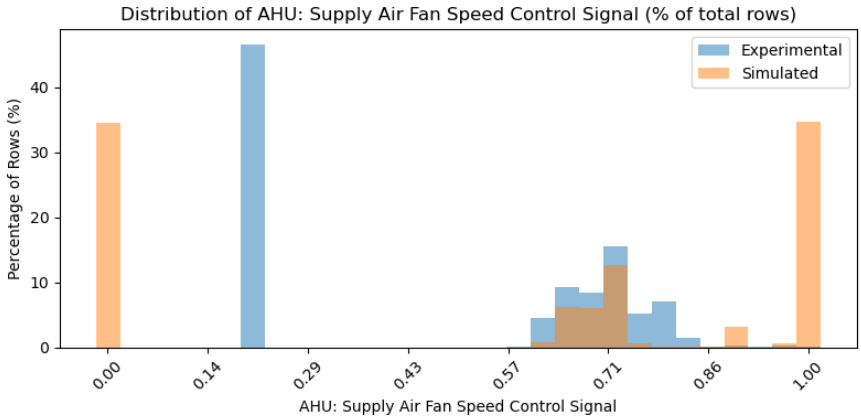

Figure 14. Supply Air Fan Speed Control Signal

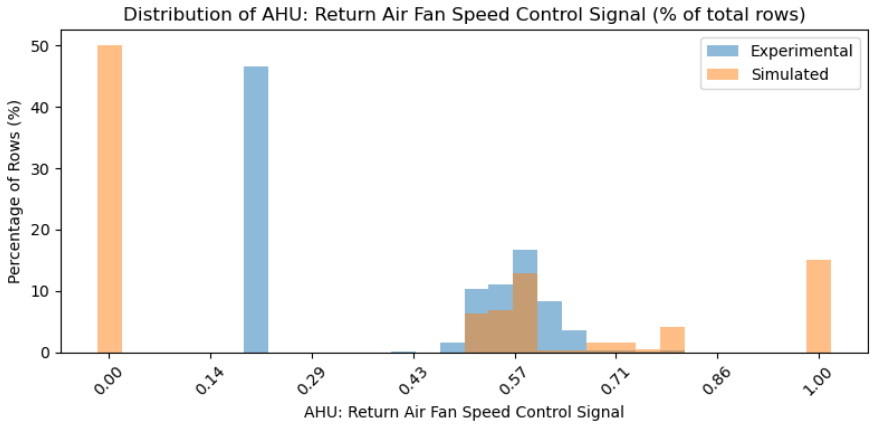

Figure 15. Return Air Fan Speed Control Signa

The figures provided above represent the air fan speed control signal during supply and return. It is observed that when using experimental data, the signal is never exactly 0, meaning that there is always some sort of air circulating in the building. It can also be seen how nearly 50% of the time the supply and return fans are almost closed, as the fans are shut off during unoccupied hours and maintained at minimum speed to meet static pressure and temperature setpoints. Moreover, it can be observed how during the occupied mode the AHU fan speed varies smoothly in a mid-range band, the values being higher in the return signal than in the supply signal. Furthermore, as mentioned in other figures the simulated data provides different scenarios, in this case approximately 70% of the time the speed control signal is off or fully on, revealing a lack of proportional control in the model.

Furthermore, to start understanding the relationship between the variables and how the sensors interact with each other, two heatmaps will be provided above. With the heatmap, it can be observed the Pearson correlation coefficients for every pair of features. The warmer tones represent a strong positive correlation between the pair of variables, while the cooler tones indicate a negative relationship. These correlation maps give a broad idea and help to identify clusters, highly correlated sensors or very features that are very independent. This idea can guide the user to choose dimensionality reduction and feature selection, which are steps that are going to be developed on the project.

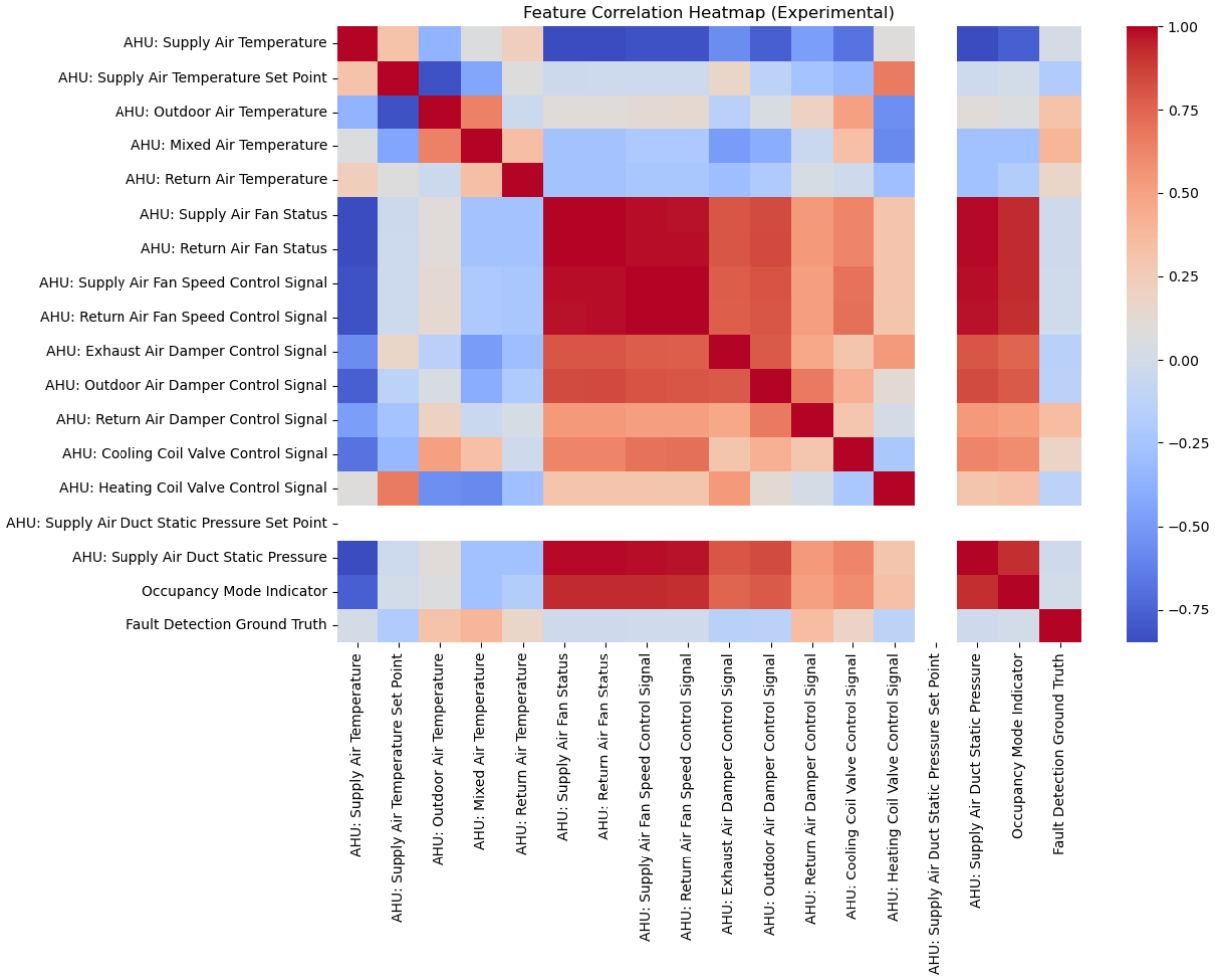

Figure 16. Feature Correlation Heatmap (Experimental)

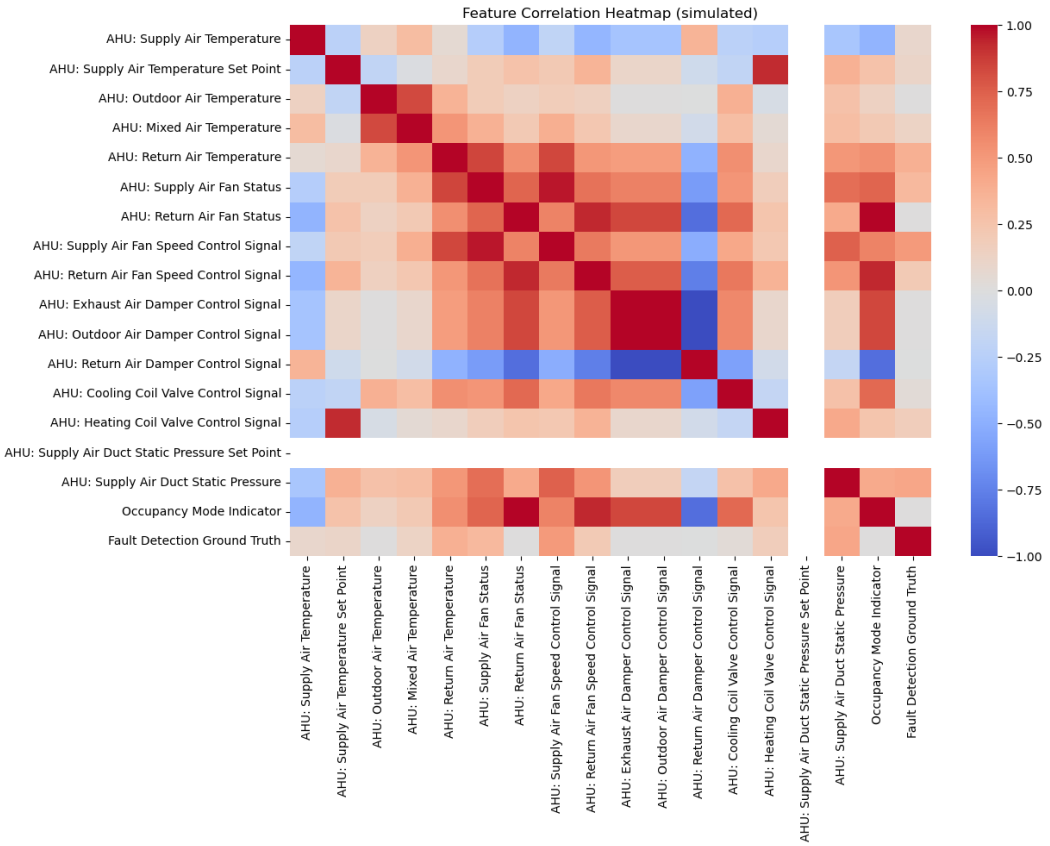

Figure 17. Feature Correlation Heatmap (Simulated)

Although the results of these heatmaps are not presented in this chapter it will be presented in the following chapter, it can be observed how in both datasets there are many correlated variables.

5.5 PRINCIPAL COMPONENT ANALYSIS (PCA)

It has been developed a Principal Component Analysis (PCA). PCA is a machine learning method used with large data sets to get dimensionality reduction, while still maintaining significant patterns and trends. Although, this technique is not necessarily useful to predict which variables give the most information to predict the faults (fault detection ground), it gives a wide idea of the data, and the amount of variance provided with several components.

It is going to be provided a graph and table for each experimental and simulated dataset. The graph shows the cumulative explained variance per number of principal components, whereas the table represents which are the variables that contribute the most to each principal component.

Moreover, it has also been calculated some of the main metrics (accuracy, precision, F1 Score and Recall), which formulas are the following:

$$1.\ Accuracy = \frac{TP + TN}{TP + FN + TN + FP}$$

$$2.\ Precision = \frac{TP}{TP+FP}$$

$$3.\ Recall = \frac{TP}{TP+FN}$$

$$4.\ F1\ Score = \frac{2*Precision*Recall}{Precision+Recall}$$

In order to calculate the different metrics, it is necessary for the reader to understand what each element in the formula accounts for:

Confusion Matrix

	Actually Positive (1)	Actually Negative (0)
Predicted Positive (1)	True Positives (TPs)	False Positives (FPs)
Predicted Negative (0)	False Negatives (FNs)	True Negatives (TNs)

Figure 18. Confusion Matrix

For a better understanding of what True Positives (TP), False Positives (FP), True Negatives (TN) and False Negatives (FN) are, a brief definition of each of those terms will be provided.

True Positives (TP): the model correctly predicts a positive case as a positive.

False Positives (FP): the model incorrectly predicts a negative case as a positive.

True Negative (TN): the model correctly predicts a negative case as a negative.

False Negative (FN): the model incorrectly predicts a positive case as negative.

5.6 SHAP RANKING

As an additional interpretive method, the SHAP (Shapley Additive Explanations) ranking was used to evaluate how much each sensor measurement contributed to the fault detection model's output. The SHAP ranking gives each feature a fair amount of credit for predicting a certain prediction and then averages the local attributions into a global ranking of importance. The SHAP ranking has two main functions for this study. First, a SHAP ranking helps to better understand which sensors have the greatest influence in the classifier model that has been trained. Second, we can use SHAP to ground a cost-effectiveness evaluation. Each sensor has its own cost for installation, calibration, and maintenance. If SHAP can aid with determining significance and importance of outputs, we can justify eliminating or simply not using the least impactful measurements without considerably sacrificing prediction. Whenever SHAP is added to the selection process for sensors, the resulting stream of data for monitoring can be reduced or specified for the sake of achieving the greatest fault-detection accuracy for each dollar spent.

5.7 COSTS

Smart buildings and HVAC systems are known for their high cost associated with both installation and ongoing operations. In practice, the total cost can vary a lot depending on

different factors. Other costs vary depending on the size of the building, building type, number of sensors used, or type of HVAC system used.

On the other hand, implementing control systems have different drivers that must be considered in deciding whether to implement HVAC systems and control systems. These main drivers are the following:

- Operational Benefits: As there is a modernization of the system, it would allow to make buildings easier to operate and could have other benefits such as user satisfaction or energy savings.
- Insights into operations: Building data is used to obtain real information of the building and make decisions based on this data. This data is often available remotely, therefore there is no need for a person to be in the building to assess the data.
- Cost Savings: although the installation of these systems does have a high cost associated with it, in the future costs associated with the building are going to decrease.
- Energy Savings: one of the main goals of this installation is to reduce the energy used.
- Ease of Use: After installation and maintenance, for the users of these facilities it is very easy.

Firstly, it is shown the costs associated with the installation of the different sensors. The total cost of the installation depends on both the hardware cost and the installation cost. To obtain the different costs of the sensors the RS Means provided by The Joint Institute for Strategic Energy Analysis [6] and other sources, especially to obtain the costs associated with the fan status. Although RS Means provides just the total range install price, it has been searched for the hardware costs and installation costs separately to provide a more detailed division of the costs.

Item	HW Cost Range ($)	HW Cost AVG ($)	Installation Cost Range ($)	Installation Cost Medium ($)	Total Range Install Price ($)	Total Average Install Price ($)
AHU: Supply Air Temperature	45-115	$ 80.00	420-3335	$ 1,085.00	465-3450	$ 1,165.00
AHU: Outdoor Air Temperature	45-115	$ 80.00	420-3335	$ 1,085.00	465-3450	$ 1,165.00
AHU: Mixed Air Temperature	45-115	$ 80.00	420-3335	$ 1,085.00	465-3450	$ 1,165.00
AHU: Return Air Temperature	45-115	$ 80.00	420-3335	$ 1,085.00	465-3450	$ 1,165.00
AHU: Supply Air Duct Static Pressure	145-440	$ 293.00	25-3305	$ 872.00	170-3745	$ 1,165.00
AHU: Supply Air Fan Status	50-150	$ 100.00	75-360	$ 218.00	125-410	$ 318.00
AHU: Return Air Fan Status	50-150	$ 100.00	75-360	$ 218.00	125-410	$ 318.00
AHU: Supply Air Fan Speed Control Signal	90-150	120	75-360	218	165-510	338
AHU: Return Air Fan Speed Control Signal	90-150	120	75-360	218	165-510	338
AHU: Cooling Coil Valve Control Signal	120-180	150	25-3305	872	145-3485	1022

AHU: Heating Coil Valve Control Signal	120-180	150	25-3305	872	145-3485	1022
AHU: Exhaust Air Damper Control Signal	140-200	170	420-3335	1085	560-3535	1255
AHU: Outdoor Air Damper Control Signal	140-200	170	420-3335	1085	560-3535	1255
AHU: Return Air Damper Control Signal	140-200	170	420-3335	1085	560-3535	1255

Table 3. Sensors' Prices

As observed in the table above, it is clearly shown how the highest portion of the total installation cost comes from the labor(installation) and not from the hardware. The labor costs account for approximately 85% of the total installation cost.

Furthermore, although it is not noted in the table above, there are other costs such as data storage or maintenance costs that also account for the total cost of a sensor.

Lastly and for a more detailed cost analysis, it has been used the data from the U.S. Department of Energy Building Technologies Office, were there is a section of Cost Category Percentages provided from interviewers [6]. The costs that are going to be used in chapter 6 are from the control's vendors.

Cost Category	Percentage	Notes	Source (interviewee type)
Overhead, Labor, and Profit	25%–33% of the installing vendor's invoice	The profit (4%–10%) is included in this number. The overhead includes administration such as accounting.	Building owner who previously worked as a vendor
Profit	4%–10% of the installing vendor's invoice		Building owner who previously worked as a vendor
Hardware	18%–20% 8%–10% technology hardware (sensors, controls) 10% is basic hardware (meters, relays)		Controls vendor
Installation Labor	40% of the invoice		Controls vendor

Table 4. Cost Category Percentages from Interviews. Source: [6]

Chapter 6. RESULT ANALYSIS

In this section the results from the system model/development described in the previous section will be provided.

6.1 DATA VISUALIZATION RESULTS

Firstly, and as it has been observed from heatmaps provided in Figure 16 and Figure 17, it can be observed how there are some highly correlated variables.

The following tables provide some high correlation combinations (>0.8), that could make people think which variables may lead to redundancy.

Correlation in the Experimental Data		
AHU: Supply Air Fan Speed Control Signal	AHU: Return Air Fan Speed Control Signal	0.999542
AHU: Supply Air Fan Status	AHU: Return Air Fan Status	0.997770
AHU: Supply Air Fan Speed Control Signal	Occupancy Mode Indicator	0.929148
AHU: Exhaust Air Damper Control Signal	AHU: Supply Air Duct Static Pressure	0.801037

Table 5: Correlation in the Experimental Data

Correlation in the Simulated Data

AHU: Return Air Fan Status	Occupancy Mode Indicator	1.00
AHU: Return Air Fan Speed Control Signal	Occupancy Mode Indicator	0.933621
AHU: Return Air Temperature	AHU: Supply Air Fan Status	0.851314
AHU: Outdoor Air Temperature	AHU: Mixed Air Temperature	0.833529

Table 6: Correlation in the Simulated Data

6.2 PRINCIPAL COMPONENT ANALYSIS RESULTS

The application of PCA provides valuable information about the variance explained in both datasets, as it shows how most of the variance from the dataset can be captured by just a few components, suggesting redundancy.

Firstly, a table and a graph show the cumulative explained variance and the top contributing features for each of the first five principal components.

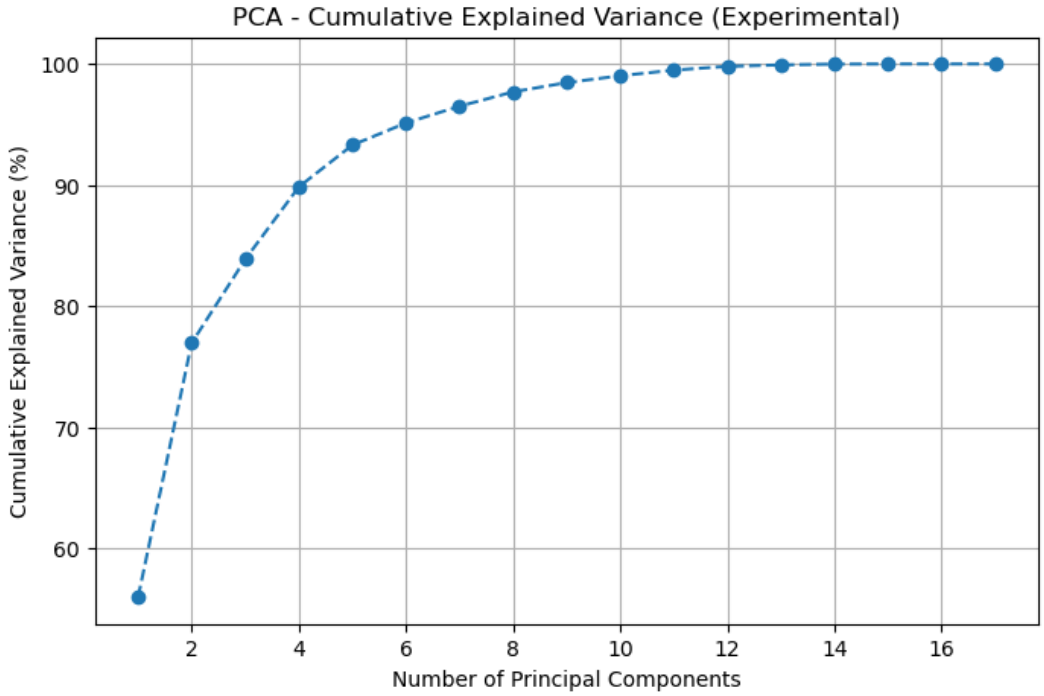

Figure 19. Cumulative Explained Variance (Experimental)

Top Contributing Features to PC1	
AHU: Return Air Fan Status	0.330458
AHU: Supply Air Fan Status	0.330140
AHU: Supply Air Duct Static Pressure	0.329918
Top Contributing Features to PC2	
AHU: Outdoor Air Temperature	0.486531
AHU: Supply Air Temperature Set Point	0.460470
AHU: Heating Coil Valve Control Signal	0.443820

Top Contributing Features to PC3	
AHU: Return Air Temperature	0.840120
AHU: Supply Air Temperature Set Point	0.339774
AHU: Mixed Air Temperature	0.249651
Top Contributing Features to PC4	
AHU: Return Air Damper Control Signal	0.646853
AHU: Cooling Valve Control Signal	0.384008
AHU: Outdoor Air Damper Control Signal	0.381482
Top Contributing Features to PC5	
AHU: Heating Coil Valve Control Signal	0.562543
AHU: Return Air Damper Control Signal	0.480568
AHU: Mixed Air Temperature	0.388777

Table 7. Top Contribution Features to PCA (Experimental)

The plot above, which represents the experimental data, shows that with four components 90% of the variance could be explained and with 6 principal components, it can capture around 95% of the total variance, meaning we can significantly reduce the number of features while keeping most useful information. Looking at the top contributing features for each component, we can see some interesting patterns. For example, the first principal component is mainly influenced by variables related to fan status and static pressure, which suggests these elements play a big role in how the system behaves overall. The second component is more about outdoor air temperature and setpoints, while the third and fourth

highlight factors like damper positions and temperature control signals. The fifth component brings attention to the heating coil valve control signal and air temperatures.

Now, the same analysis is provided but for the simulated data

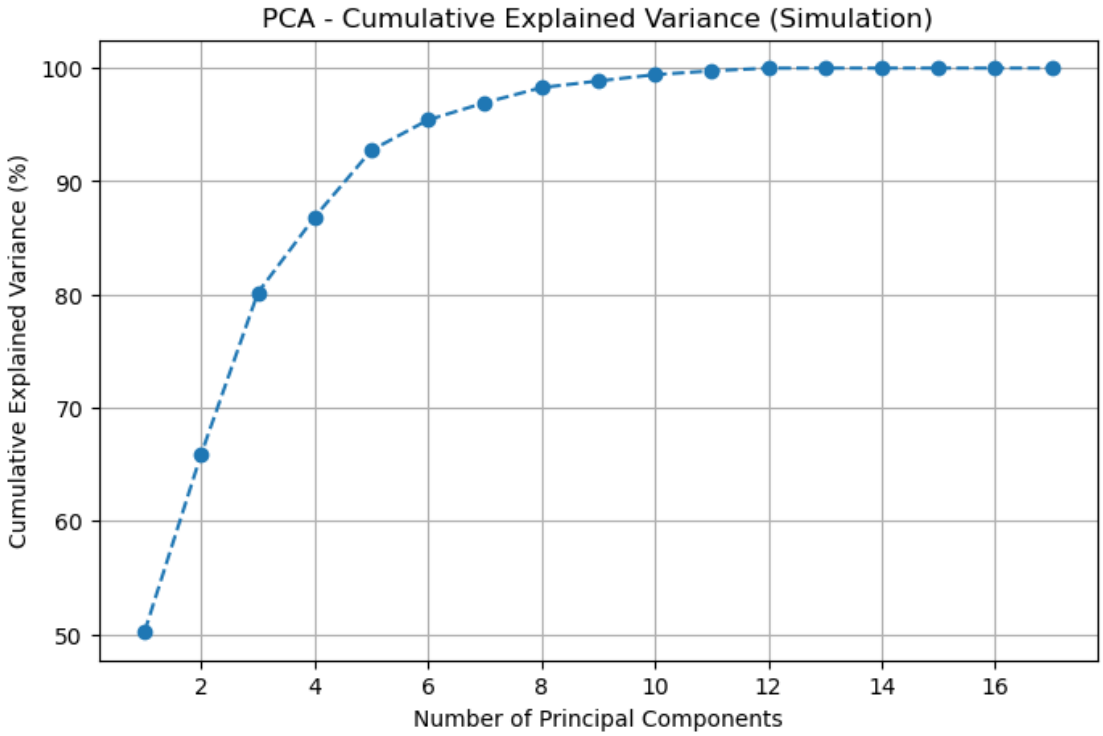

Figure 20.Cumulative Explained Variance (Simulation)

Top Contributing Features to PC1	
AHU: Return Air Fan Status	0.333678

Occupancy Mode Indicator	0.330140
AHU: Return Air Fan Speed Control Signal	0.323932
Top Contributing Features to PC2	
AHU: Mixed Air Temperature	0.498833
AHU: Outdoor Air Temperature	0.492793
AHU: Supply Air Temperature	0.336570
Top Contributing Features to PC3	
AHU: Heating Coil Valve Control Signal	0.521613
AHU: Supply Air Temperature Set Point	0.500996
AHU: Supply Air Duct Static Pressure	0.395121
Top Contributing Features to PC4	
AHU: Outdoor Air Temperature	0.443607
AHU: Supply Air Fan Speed Control Signal	0.419942
AHU: Supply Air Fan Status	0.360421
Top Contributing Features to PC5	
AHU: Supply Air Temperature	0.669041
AHU: Supply Air Duct Static Pressure	0.329470

AHU: Cooling Coil Valve Control Signal	0.288077

Table 8. Top Contribution Features to PCA (Simulation)

As it can be observed from the graph above, the simulated data set exhibits a lower explained variance compared to the experimental data set. This disparity arises from the difference in data origins. While experimental data is composed from real-world data, the simulated data set incorporates alterations and new scenarios. These new additions increase complexity, making it more difficult to explain the variance of the data set when new cases are introduced.

From the graph and table referring to the Principal Component Analysis for simulated data, it can be observed how with less than four components the explained variance is lower than in the experimental data, but as more components are introduced the cumulative explained variance improves reaching nearly a 98% of cumulative variance with eight components, similar to the experimental data.

The PCA decomposes system behavior along five main modes: the first node driven by airflow and occupancy patterns, the second reflecting temperature patterns, two capturing the control actuated by valves and dampers and third capturing cooling-specific dynamics. The PCA-focused understanding demonstrated that the vast majority of variability first comes from how air is moved and conditioned and then the next order of variability related to steps taken with adjustments to control provisioned heating and cooling.

6.3 METRICS FOR PCA

With the insights provided from PCA different metrics have been calculated in different scenarios. These different scenarios have been calculated with different Machine Learning supervised algorithms, which were Decision Trees, Random Forest and XGBoost. These

algorithms bring unique strengths to analyzing and predicting outcomes based on labeled data.

Firstly, the variables which contribute the most to the first four PCs have been used.

	Accuracy (%)	**Precision (%)**	**Recall (%)**	**F1 Score (%)**
Decision Tree	98.72	95.19	95.08	95.14
Random Forest	98.86	95.35	96.02	95.68
XGBoost	98.90	95.58	96.14	95.86

Table 9. Metrics using the best feature for the first four PCs (Experimental)

If we use simulated data, with the features that contribute the most to the first four PCA to simulated data, we obtain the following metrics.

	Accuracy (%)	**Precision (%)**	**Recall (%)**	**F1 Score (%)**
Decision Tree	79.88	83.76	74.15	78.66
Random Forest	79.47	81.90	75.69	78.67
XGBoost	83.93	94.68	71.91	81.74

Table 10. Metrics using the best feature for the four fisrt PCs (Simulation)

Additionally, when using the variables with contributes the most to the experimental data in the simulated data, the main metrics are the following:

	Accuracy (%)	**Precision (%)**	**Recall (%)**	**F1 Score (%)**

Decision Tree	81.09	82.75	78.57	80.61
Random Forest	81.47	82.35	80.14	81.23
XGBoost	84.49	89.26	78.43	83.49

Table 11. Metrics using the best feature for the first four PCs in experimental data but using simulation data)

As shown in Table 9, all three models achieve excellent results with accuracies of 98.7 to 98.9 % and F1 scores of 95.1 to 95.9 % (with XGBoost roughly better), when applying the top four PCs from the experimental data. However, Table 10 shows that when we train and test only on the PCs from the simulated data, accuracy falls to 79.5–83.9 % and F1 to 78.7–81.7 %, indicating that the simulated datd fails to obtain important predictive variation. Table 11 shows that simply replacing the experimental PCs for the simulated data recovers much of that lost ground (81.1–84.5 % accuracy and 80.6–83.5 % F1 across models). Taking altogether, these results demonstrate that principal components computed from real measurements capture better the overall variance of the dataset.

Furthermore, the resolution of the data has been changed. The original dataset took a shot of the system every minute, but it has been tried to reduce the resolution to take a snapshot of the data every fifteen minutes and each hour to observe how the metrics change. Although data storage is not the principal cost of an HVAC system, it helps to reduce costs. The metrics that are going to be shown are the accuracy and F1 Score as they reveal the most.

	Accuracy (%)		F1 Score (%)	
	Resolution:15’	Resolution:1h	Resolution:15’	Resolution:1h

Decision Tree	97.92	95.37	91.09	83.87
Random Forest	97.22	96.30	88.46	86.67
XGBoost	96.76	95.37	86.54	83.87

Table 12. Metrics using the best feature for the first four PCs with different resolution (Experimental)

	Accuracy (%)		F1 Score (%)	
	Resolution:15'	Resolution:1h	Resolution:15'	Resolution:1h
Decision Tree	67.29	60.11	66.58	60.73
Random Forest	69.96	63.83	69.05	64.58
XGBoost	76.50	63.83	73.81	64.21

Table 13. Metrics using the best features for the first four PCs with different resolutions (Simulation)

	Accuracy (%)		F1 Score (%)	
	Resolution:15'	Resolution:1h	Resolution:15'	Resolution:1h
Decision Tree	70.89	65.96	71.16	67.68
Random Forest	72.90	69.15	72.82	69.79
XGBoost	74.23	68.09	74.16	69.07

Table 14. Metrics using the best features for the first four PCs in experimental data but using different resolutions with simulation data

As it can be observed from the three table provided above, across our experimental dataset, as the sampling interval changes from one minute to fifteen minute to one hour, it is observed a bit of reduction in performance, with accuracy falling from around 98%–99% at one minute to around 95%–96% at one hour, and F1 scores falling from approximately 91% at fifteen minutes to nearly 84% at one hour (Table 12). In contrast, the simulated data are subject to much noticeable reduction (Table 13): accuracy falls from ~76% at 1 minute resolution to around 60% at one hour, and F1 falls from around 78% to approximately 64%. When we then assess our experimental PCs with the simulated dataset (Table 14), we notice the effects of lower resolution are somewhat reduced as accuracy drops around 3-5% from a 15 mins resolution to an hour resolution but is still many points below the original minute-by-minute results (drops around 10% from 1 min resolution to 15 mins resolution).

These results show how capturing more data points gives better metrics in knowing the target variable (Fault Detection Ground Truth). It can be observed how capturing real-world data tolerate sampling is much better than with simulation data.

Furthermore, given the best feature for each of the first four PCA in experimental data and with the supervised machine learning techniques used: Decision Trees (DT), Random Forest (RF) and XGBoost (XGB), it is provided the best combinations.

Model Used	Features	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
XGB	All	98.90	95.58	96.14	95.86
RF	All	98.86	95.35	96.02	95.68
DT	All	98.72	95.19	95.08	95.14
XGB	OAT,RAT,RADCS	98.69	94.45	95.67	95.06

RF	OAT,RAT,RADCS	98.83	95.13	96.02	95.57
DT	OAT,RAT,RADCS	98.58	94.41	94.85	94.63
XGB	RAFS, OAT, RADCS	96.82	84.91	92.27	88.44

Table 15. Different models and features metrics combination

It is important to note the abbreviations used in the Features Column:

- OAT: AHU Outdoor Air Temperature
- RAT: AHU Outdoor Air Temperature
- RAFS: AHU Return Air Fan Status
- RADCS: AHU Return Air Damper Control Signal

6.4 SHAP ANALYSIS RESULTS

In this section, it is going to be shown the SHAP ranking for the experimental data and the metrics using different combination of variables.

The illustration provided below shows the SHAP ranking for the experimental data.

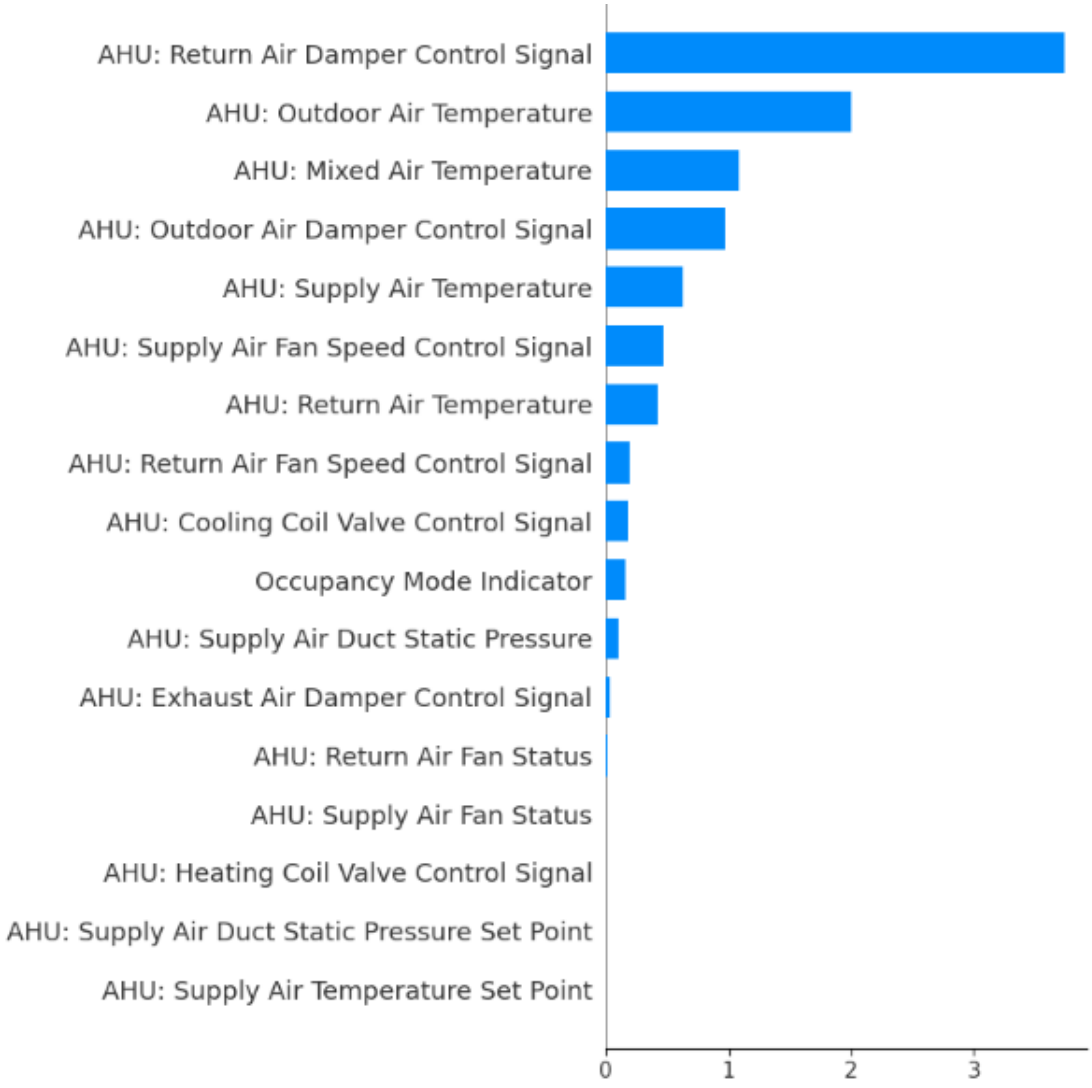

Figure 21. SHAP Ranking for Experimental Data

In Figure 21, the horizontal bars of average SHAP values highlight exactly which sensors the fault-detection model depends on for decision making. The first bar returns air damper control signal, which clearly indicates that small adjustments to damper positioning give idea an idea of a system issue or a warning of emerging problems. This is closely followed by outdoor air temperature and mixed air temperature. Together, they emphasize that the model is particularly sensitive to changes in outdoor air conditions and the mixing of incoming fresh and recirculating air. When outdoor air conditions change suddenly such as an unexpected hot heatwave or an inflow of cold air, the model tends to notice these changes immediately as those types of shifts usually precede comfort issues or control loop instability.

In contrast, some of the variables that influence the least to capture whether the system's behavior is correct or not, are set points referring to the air status, the pressure or the heating coil valve control signal.

This is particularly important for practical reasons: it shows that beyond the few measurements at the top of this SHAP chart, several sensors could be identified and either retired or deprioritized with little impact on the model's detection performance. Importantly, all this information comes from the experimental dataset, selected rather than simulated output because it gives real life behavior: the calibration drifts, the sensor noise and the precise relationships between components that a simple simulator cannot reproduce. This means, or said differently, that this reflects the true dynamics of the AHU in operation and identifies the measurements that will provide practitioners with the most robust diagnostic information when they need it in practice.

Now, the metrics using these top four features will be shown.

	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Decision Tree	98.83	95.13	96.02	95.57
Random Forest	98.97	95.18	97.07	96.12

XGBoost	99.09	95.95	97.19	96.57

Table 16. Metrics using the best four features for SHAP Ranking (Experimental)

	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Decision Tree	79.58	83.56	73.66	78.30
Random Forest	79.35	82.04	75.19	78.46
XGBoost	83.60	93.87	71.91	81.44

Table 17. Metrics using the best four features for Experimental SHAP Ranking (Simulation)

As it is can be observed from Table 16 and Table 17, metrics have a better performance when using selected variables from the SHAP ranking than the Principal Component Analysis. For instance, in real data, the XGBoost drives the accuracy up to 99.09 % and F1 to just under 96.57 % compared to 98.86 % accuracy and 95.68 % F1 when using the top four component scores. On the other hand, if the simulated data is selected using the best features for experimental data the difference is not that pronounces going the XGBoost from under 83.49 % F1 to 83.60 % once it is presented with just a few key sensors identified using SHAP.

Although, it has been mentioned above, it is important to understand the reason why the best features for the simulated data have not been selected, although their metrics are better than the metrics if using the experimental SHAP ranking. Some of the main reasons include:

- Real sensors include noise, movements or other non-expected things that clean simulated data does not reproduce.

- Models may be overfit to simulator's idealized patterns.
- Experimental fault labels could be delayed.
- Actual operating conditions do not expect unmodeled variability but can happen.

Moreover, the behavior of the F1 score is going to be plotted under the scenario of being used the 3,4 or 5 SHAP ranked features on the experimental dataset. The metric used has been the F1 score as it is a harmonic mean of precision and recall. This metric provides useful insights as it takes a lot into account the false alarms which carry significant costs.

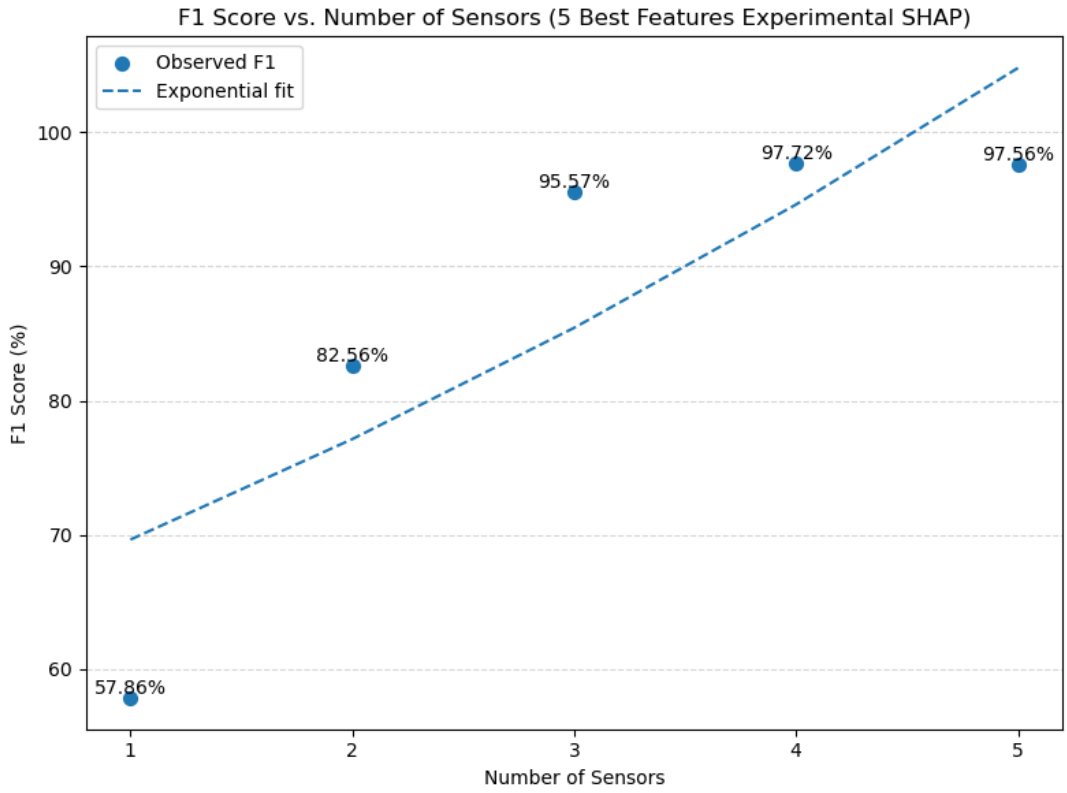

Figure 22. F1 Score vs. Number of Sensors (5 Best Features Experimental SHAP)

In the chart above, it is evident that adding a fourth SHAP ranked feature), increases the F1 score substantially (95.57 % to 97.72 %). However, adding a fifth feature decreases performance slightly (97.72 % to 97.56 %). It is then suggested to reduce the features to the best four SHAP ranking features, as it has reached the elbow and it keeps the system and model as simple as possible.

6.5 COSTS RESULTS

In this section it is going to be described the cost analysis and observing how much the cost is reduced in different scenarios.

Firstly, it is shown the total installation cost depending on the number of features selected. It is studied, as in the SHAP ranking analysis section, how does the installation cost varies if the top three, four or five SHAP ranking features are used, as well as if there was no feature reduction (full set).

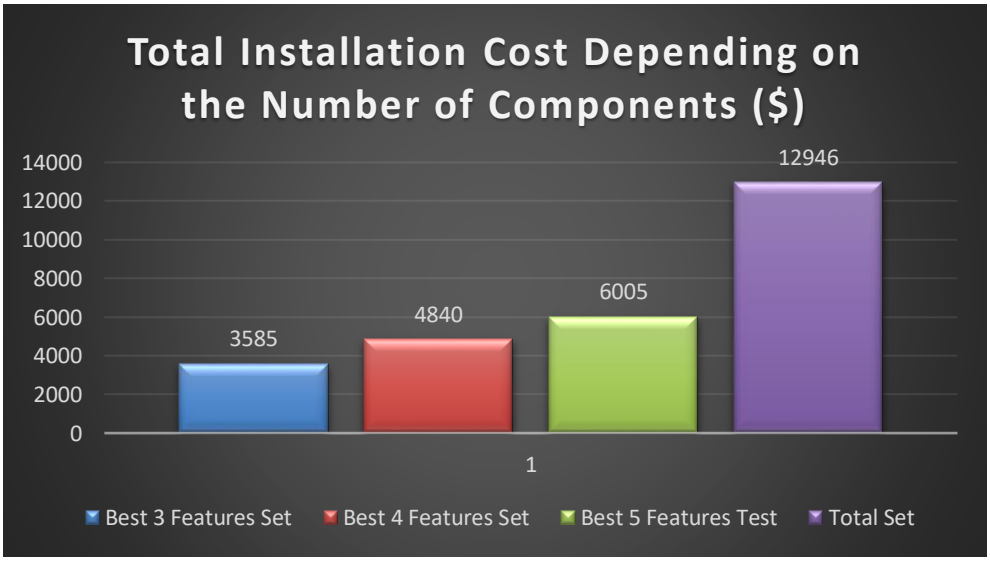

Figure 23. Total Installation Cost Depending on the Number of Components without Overhead and Profits Costs ($)

As can be observed in Figure 23, the cost difference between four or five variables is 1165$, while the F1Score drops. On the opposite side, it is decided that for a 2.15% F1 score it is worth the investment of an additional 1255$ per Air Handling Unit (AHU). In this building, as there are three AHUs, it would be 3765$.

Now, the cost analysis will be provided for the four sensors set and the full set. The hardware cost analysis, labor analysis and total installation analysis are provided, which is the sum of the hardware and labor cost. Once again, this graph represents just an AHU, therefore the cost would triple.

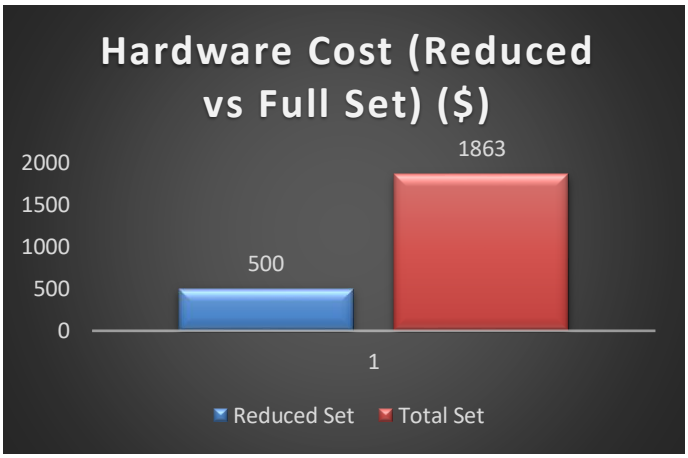

Figure 24. Hardware Cost (Reduced vs Full Set)

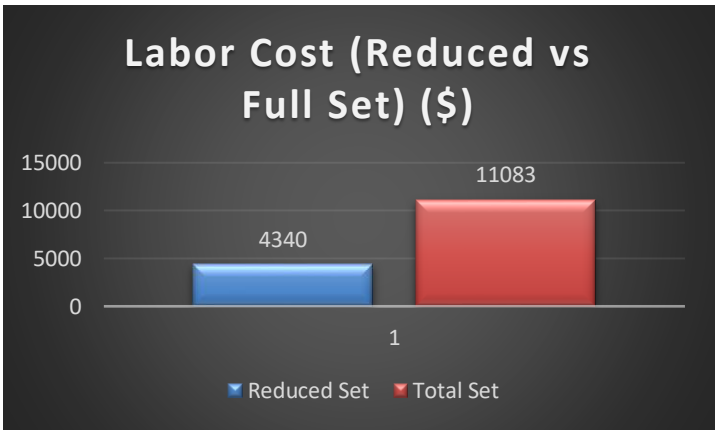

Figure 25. Labor Cost (Reduced vs Full Set)

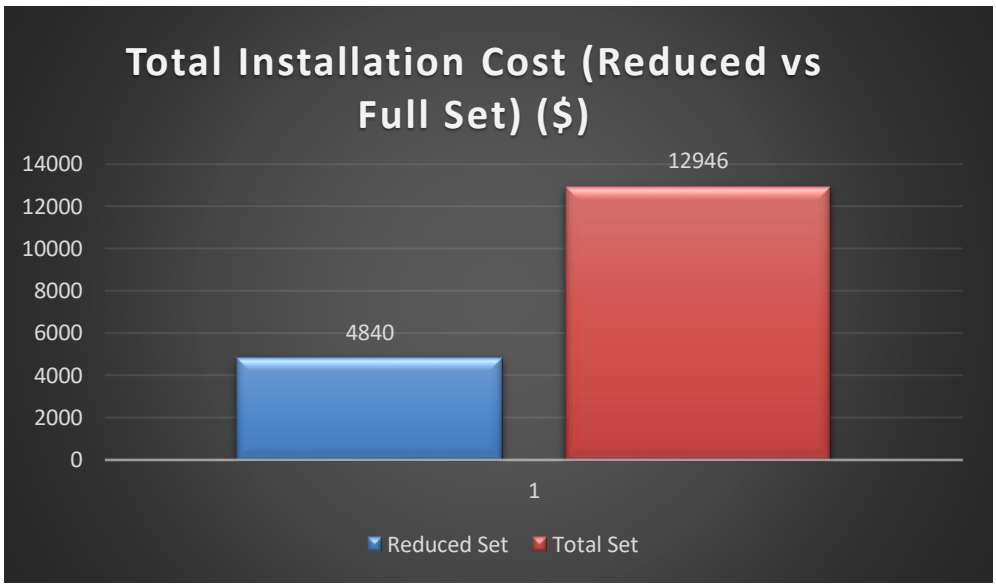

Figure 26. Installation Cost (Reduced vs Full Set) without Overhead and Profits Costs ($)

As it can be observed from the three figures provided above, under this building setup the cost of AHU installation without overhead and profits is of 12946$ per AHU. Therefore, as this building is composed of three AHUs and imagining each AHU is composed of the same sensors, the cost would be 38838$.

On the other hand, if the set was composed only of the top 4 SHAP ranking variables in the experimental data, the cost per AHU (without overhead and profits) would be 4840$, being the total cost of the three AHUs of 14520$, which is 37% of the cost if a full set were used.

Moreover, although the labor and hardware prices vary depending on the sensor, it can be observed how between 10-17% of the total installation cost accounts for the hardware and the remaining 83-90% for the labor.

Next and focusing only on the four variables set and the total set, it is provided a breakdown of their costs into hardware, labor and overhead + profit. For calculating this data, it has been used Table 4 cost breakdown which source is controls vendors.

As for the calculation of the total installation cost, it has been used the medium values, it has been considered for the overhead+ profit to take also into account medium values, where the overheads + profit accounts for 40%. Therefore:

- Raw base cost: Labor (Figure 25) + HW Cost (Figure 24) = Installation Cost (Figure 26)
- Invoice: Raw Base Cost / (1- 40% (Overhead + Profit))
- Overhead + Profit = Invoice * 0,4

The total breakdown of the cost for one Air Handling Unit (AHU) would be:

	Hardware ($)	Labor ($)	Overhead + Profits ($)	Total Invoice ($)
Reduced Set of Variables	500	4340	3227	8067
Full Set	1863	11083	8630	21577

Table 18. Breakdown of Costs for One Air Handling Unit (AHU)

And in consequence, for this system in this building the breakdown for the three AHUs would be:

	Hardware ($)	Labor ($)	Overhead + Profits ($)	Total Invoice ($)

Reduced Set of Variables	1500	13020	9681	24201
Full Set	5589	33249	8630	64731

Table 19. Breakdown of Costs for the Complete Building

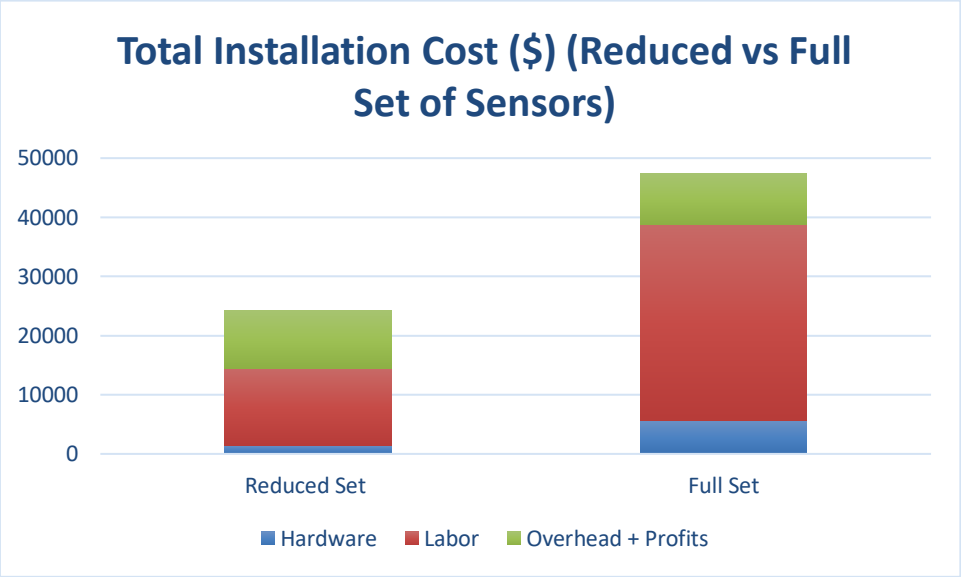

Figure 27. Total Installation Cost ($) (Reduced vs Full Set of Sensors)

As can be observed in the graph provided above, there could be a significant cost expenditure if the sensors used were not all the ones included in the data set but were used the best four SHAP ranking sensors.

Chapter 7. CONCLUSIONS AND FUTURE WORK

This project has conducted a complete study of the tradeoffs in sensor availability and fault detection capability in Heating, Ventilation, and Air Conditioning (HVAC) systems. The objectives clearly defined in this project have been met, highlighting the feasibility of selecting a reduced number of sensors while only minimally sacrificing the reliability of anomaly detection in HVAC. With this project it is given a meaningful contribution at the technical and economical levels in the field of smart buildings.

7.1 PROJECT CONCLUSION

Firstly, it is used different supervised machine learning techniques which include Decision Trees, Random Forest and XGBoost models to test a strong fault detection system. It has been very relevant to choose these algorithms as they show a great ability to learn from labeled datasets and accurately predict the presence of anomalies in the HVAC system, thanks to the labelled variable (Fault Detection Ground Truth). The datasets used experimental and simulated data. The diverse datasets allowed us to test and validate models' applicability over a vast array of conditions, representing real-world scenarios as well as forced scenarios so that we could better understand system behavior.

An important part of the project reflects the application and the analysis of the Principal Component Analysis (PCA) and SHAP (Shapley Additive Explanations) to find out the most informative features. PCA analysis showed that with a reduced number of sensors it could be explained most of the total data variance (with 5 variables it could be shown approximately 95% of the variance). In the case of the data, the experimental data components captured more of a generalizable discriminative structure than with that of the simulated data. This shows us the importance of developing these optimizations based on real world behavior.

In addition to PCA, it is used the SHAP ranking to provide a better understanding of the contribution of each sensor to the model prediction. The sensors considered the most impactful included signals such as the return air damper control, outdoor air temperature, and mixed air temperature. Therefore, the results offered a basis for eliminating or deprioritizing insignificant sensors and still have a valid detection process. The performance metrics were a bit better than when using the PCA feature selection, so it can be valid as a good sensor reduction technique.

During the project it has been quantified the direct relationship between the number of sensors used and the model's performance. It has been demonstrated that an acceptable level of accuracy and reliability can be obtained in detecting anomalies in the systems with an intelligent subset of sensors. It has been tested that a reduction in the temporal resolution affects reasonable the performance of the system. Moreover, it has been proved that experimental data shows a more realistic behavior of the system as it considers calibrations or shifts.

Lastly, it has been conducted an economic impact of sensor reduction. It is shown how a reduction in the number of sensors can be traduced to a lot of savings. These savings are not just the initial payments associated with the installation and the hardware but also in future expenses such as maintenance or data storage. The most amount of the sensor installation cost comes from labor, therefore it is fundamental to know which sensors are the best to predict anomalies. The project provides a basis for developing better sustainability and lower cost HVAC system designs by identifying sensor configurations that achieved high accuracy and low cost.

To sum up, the main goals planned at the beginning of the project have been achieved. The objectives include the development of an accurate fault detection system, which evaluates real and simulated HVAC data. As the main goal is to reduce the cost of the HVAC system and evaluate the economic impact, it has been applied dimensionality reduction and feature selection techniques. The most useful insight has been that sensor reduction is feasible and

beneficial as the metrics do not reduce much, and therefore, a reduction in the HVAC system is achieved.

7.2 FUTURE WORK

With the development of this project, it is manifested there is future work that can be done.

1. Generalization and Flexibility to Different Systems and Buildings: An important part for future work is attributed to testing and adapting this work methodology to other HVAC systems, types of building or sizes as the features and metrics might change depending on the circumstances. The findings shown in this paper reflect the results for a specific dataset from a relatively small commercial building. Exploring the transferability and stability of reduced sensor fault detection algorithm, and sensor selection approaches in different operational profiles would represent an initial feature and would help to develop a large-scale implementation. Such an approach would require developing a more diverse database and applying transfer learning techniques.
2. Integration of Multi-Source and Contextual Data: Smart buildings generate a very high volume of data from different sources, not just HVAC sensor data. Knowing what other sensors are mentioned would allow us to have wider, more precise and adaptive models of the building.
3. More accurate Fault Detection and Diagnosis: It is recommended for future dataset to provide different classification and granular diagnosis of specific fault types such as stuck damper, error in the fan or sensor drift. This would require a more detailed dataset with labels for each specific fault type. This could allow professionals to identify the cause of the problem quicker and anticipate maintenance efforts.
4. Predictive and Prescriptive Models: New models could be built to predict sensor faults before they happen and even suggest the best action to take for more efficient building management.

5. Deeper Cost Analysis: Future works in this field could obtain deeper research of the cost analysis, looking for the false positive and false negative cost, and what would happen if these faults were undetected.
6. Multi-Objective Optimization of Cost and Performance: Suggest a multi-objective optimization and performance not only talking about accuracy and potential cost savings, but including other factors such as energy usage, system life span and occupant comfort. This may include the development of optimization algorithms in which objectives could consider values of tradeoffs space to find the best compromise solutions.
7. Reduction of Data in Other Areas with High Data Volume: Although this project is for HVAC systems, the methodology and conclusions could be applied to different areas with high volume of data, where data can be reduced while optimizing efficiency and errors detection. Below are going to be shown sectors with high volumes of data:
    - Smart cities: Modern cities use lots of sensors such as cameras, quality of air sensors, smart street lighting networks and other sensors for controlling the water and energy supply systems. Decrementing sensor redundancy could lead to a more efficient and sustainable city.
    - Manufacturing and Industrial Automation (Industry 4.0): Within manufacturing environments, there are essentially an infinite number of connected sensors (temperature, pressure, vibration, current, etc.) throughout the machines and production lines. By applying similar techniques, predictive maintenance of machinery can be more optimized, overtime reducing downtime, improving product quality and overall operating costs, as it will identify the most critical sensors for monitoring equipment health.
    - Transportation Systems (Autonomous Vehicles, Railways): New automobiles and railways involve a wide variety of sensors for navigation, performance monitoring, and safety. Intelligent sensor selection and optimization are critical to ultimately ensure reliability and reduce the complexity and cost of the systems.

- Social Media Data: Social media data is all over the net, including texts, images, interactions, etc. By creating an algorithm that detects patterns to avoid spam, misinformation and still be able to develop a targeted advertisement, it can help to reduce computation costs.

To summarize, this project has established a solid starting point in the space of fault detection systems in HVAC. Next steps not only involve further establishing these technical and economic factors in this area but also transferring these concepts and approaches to other data-centric fields where efficiency, sustainability, and reliability are also important.

Chapter 8. BIBLIOGRAPHY

1. S. E. Mills, M. J. Winiarski, and M. A. Friedman, "Analysis of Automated Fault Detection and Diagnostics (AFDD) Records as an Indicator of HVAC Fault Prevalence: Methodology and Preliminary Results," Pacific Northwest National Laboratory, Richland, WA, PNNL-31804, 2021. [Online]. Available: http://eta-publications.lbl.gov/sites/default/files/analysis_of_automated_fault_detection_and_diagnostics_records.pdf
2. X. Zhou, P. Haves, and S. E. Lee, "Quantifying energy savings from fault detection and diagnosis," *Energy Engineering*, vol. 102, no. 5, pp. 6–26, 2005. doi: 10.1080/10789669.2005.10391123. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/10789669.2005.10391123
3. I. Matetić, I. Štajduhar, I. Wolf, and S. Ljubic, "A review of data-driven approaches and techniques for fault detection and diagnosis in HVAC systems," *Sensors*, vol. 23, no. 1, p. 1, 2023. [Online]. Available: https://doi.org/10.3390/s23010001
4. J. Chen, L. Zhang, Y. Li, Y. Shi, X. Gao, and Y. Hu, "A review of computing-based automated fault detection and diagnosis of heating, ventilation and air conditioning systems," *Renewable and Sustainable Energy Reviews*, vol. 161, p. 112395, 2022. [Online]. Available: https://doi.org/10.1016/j.rser.2022.112395
5. Z. Chen *et al.*, "A review of data-driven fault detection and diagnostics for building HVAC systems," *Applied Energy*, vol. 339, p. 121030, Jun. 2023. [Online]. Available: https://doi.org/10.1016/j.apenergy.2023.121030
6. K. Trenbath, R. Meyer, K. Woldekidan, K. Maisha, and M. Harris, *Commercial Building Sensors and Controls Systems – Barriers, Drivers, and Costs*, National Renewable Energy Laboratory, Tech. Rep. NREL/TP-6A50-82117, Aug. 2022. [Online]. Available: https://www.nrel.gov/docs/fy22osti/82117.pdf
7. N. Daurenbayeva, L. Atymtayeva, A. Nurlanuly, A. Bykov, B. Akhmetov, G. Shuitenov, and U. Turusbekova, "A machine learning approach to microclimate

monitoring and fault detection," *Appl. Math. Inf. Sci.*, vol. 19, no. 2, pp. 327–334, Mar. 2025. [Online]. Available: https://doi.org/10.18576/amis/190209

8. L. Nie, R. Wu, Y. Ren, and M. Tan, "Research on fault diagnosis of HVAC systems based on the ReliefF-RFECV-SVM combined model," *Actuators*, vol. 12, no. 6, p. 242, Jun. 2023. [Online]. Available: https://doi.org/10.3390/act12060242
9. Y. Chen, A. Ebrahimifakhar, and D. P. Yuill, "Using machine learning for feature selection in automated fault detection and diagnosis of split system air conditioners," in *Proc. Int. Refrigeration and Air Conditioning Conf.*, Purdue University, West Lafayette, IN, USA, 2022. [Online]. Available: https://docs.lib.purdue.edu/iracc/2403/
10. J. Granderson, G. Lin, A. Harding, P. Im, and Y. Chen, "Dataset for building fault detection and diagnostics algorithm creation and performance testing, Version 3," *figshare*, 2020. [Online]. Available: https://doi.org/10.6084/m9.figshare.11752740.v3
11. J. Wen, S. Pourarian, X. Yang, and X. Li, *Tools for Evaluating Fault Detection and Diagnostic Methods for HVAC Secondary Systems of a Net Zero Building*, Final Report, NIST 10D243, Drexel University, prepared for the U.S. National Institute of Standards and Technology, Gaithersburg, MD, USA, Jun. 2015. [Online]. Available: https://www.researchgate.net/profile/Jin-Wen-12/publication/337946061_NIST10D243_FinalReport_-Tools_for_Evaluating_Fault_Detection_and_Diagnostic_Methods_for_HVAC_Secondary_Systems_of_a_Net_Zero_Building/links/5df73f11299bf10bc35f1163/NIST10D243-FinalReport-Tools-for-Evaluating-Fault-Detection-and-Diagnostic-Methods-for-HVAC-Secondary-Systems-of-a-Net-Zero-Building.pdf
12. Home – Lawrence Berkeley National Laboratory
13. https://squarehvac.com/parts-of-hvac-system/
14. https://globalsummitryproject.com/focusing-on-the-future-where-are-we-on-the-sustainable-development-goals-sdgs-and-some-other-things/
15. R. I. Jony, R. I. Rony, M. Rahman, and A. Rahat, "Big Data Characteristics, Value Chain and Challenges," in Proc. 1st Int. Conf. on Advanced Information and Communication Technology (ICAICT), Chittagong, Bangladesh, May 2016. [Online]. Available:

https://www.researchgate.net/publication/311323742_Big_Data_Characteristics_Value_Chain_and_Challenges

ANEXX I: PROJECT ALIGNMENT WITH THE SDGS

This project is primary aligned with several United Nations Sustainable Development Goals (SDGs) by contributing with buildings and technical efficiency, sustainable infrastructure and energy optimization in buildings. The SDGs represent a global plan to achieve a better and more sustainable future, and this project contributes directly to several of the goals.

Firstly, and to enhance the readers' understanding, I will present and therefore plot, which are the 17 different SDGs:

Figure 28. Sustainable Development Goals (SDGs) [14]

Below is an explanation of how this project relates to specific SDGs:

SDG 9: Industry, Innovation, and Infrastructure:

The project helps to produce more resilient and efficient systems by improving fault detection capabilities, resulting in a stronger infrastructure that can resist and recover from potential breakdowns.

SDG 11: Sustainable Cities and Communities:

By improving fault detection processes in buildings, the project helps to make buildings (and therefore cities) smarter and more sustainable, ensuring better management of resources in buildings.

SDG 12: Responsible Consumption and Production:

This project promotes resource efficiency by reducing unnecessary sensor usage while ensuring system reliability.

SDG 13: Climate Action

This project promotes the optimization of energy consumption in smart buildings through more efficient fault detection systems and the reduction of sensors. HVAC systems are among the highest energy consuming components buildings. This project helps to reduce unnecessary energy usage, enhance operational performance and therefore support decarbonization efforts.

SDG 17: Partnership for Goals

This project encourages collaboration between academic institutions, industry experts and data communities. With the use of shared data (in this case LBNL data), research exchanges, interdisciplinary cooperation, the project guides collective progress toward more efficient and sustainable technological solutions.

Additionally, while the most direct impact of this project aligns with the goals mentioned above, its larger contribution is to advance the sustainable engineering approach. It proves that efficiency, interpretability and long-term effects should be considered while designing solutions. This project serves as a step towards a more environmental consciousness when talking about engineering methods.

Furthermore, the methodology and findings of this project can serve future work in other domains, such as water systems or energy management and new goals could be also under scope.

ANNEX II

Some of the code developed will be provided above.

PCA:

```
##Experimental
# 1. Separate features and target
featuresexp = dataexp.drop(columns=['Fault Detection Ground Truth', 'Datetime'])
target = dataexp['Fault Detection Ground Truth']

# 2.Standardize the features
scaler = StandardScaler()
X_scaled = scaler.fit_transform(featuresexp)

# 3.Perform PCA
pca = PCA()
X_pca = pca.fit_transform(X_scaled)

# 6. Explained variance (to see how many components explain most of the variance)
explained_varianceexp = pca.explained_variance_ratio_
cumulative_varianceexp = np.cumsum(explained_varianceexp)

plt.figure(figsize=(8,5))
plt.plot(range(1, len(explained_varianceexp) + 1), cumulative_varianceexp*100,
marker='o', linestyle='--')
plt.xlabel('Number of Principal Components')
plt.ylabel('Cumulative Explained Variance (%)')
plt.title('PCA - Cumulative Explained Variance (Experimental)')
plt.grid(True)
plt.show()

# 7. PCA Loadings (Feature contribution to each PC)
loadingsexp = pd.DataFrame(pca.components_.T,
                        columns=[f'PC{i+1}' for i in
range(len(pca.components_))],
                        index=featuresexp.columns)

# 8. Check top contributing features for PC1 to PC5
top_features_pc1 = loadingsexp['PC1'].abs().sort_values(ascending=False).head(5)
top_features_pc2 = loadingsexp['PC2'].abs().sort_values(ascending=False).head(5)
top_features_pc3 = loadingsexp['PC3'].abs().sort_values(ascending=False).head(5)
top_features_pc4 = loadingsexp['PC4'].abs().sort_values(ascending=False).head(5)
top_features_pc5 = loadingsexp['PC5'].abs().sort_values(ascending=False).head(5)

print("\nTop Contributing Features to PC1:")
print(top_features_pc1)
print("\nTop Contributing Features to PC2:")
```

```
print(top_features_pc2)
print("\nTop Contributing Features to PC3:")
print(top_features_pc3)
print("\nTop Contributing Features to PC4:")
print(top_features_pc4)
print("\nTop Contributing Features to PC5:")
print(top_features_pc5)
```

Function of evaluation of metrics and for running a fault detection analysis

```
import numpy as np
from sklearn.model_selection import train_test_split
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import RandomForestClassifier
from xgboost import XGBClassifier
from sklearn.metrics import accuracy_score, precision_score, recall_score,
f1_score
from sklearn.preprocessing import StandardScaler

# Evaluation of the different metrics
def evaluate(y_true, y_pred):
    return {
        "Accuracy": round(accuracy_score(y_true, y_pred) * 100, 2),
        "Precision": round(precision_score(y_true, y_pred, zero_division=0) *
100, 2),
        "Recall": round(recall_score(y_true, y_pred, zero_division=0) * 100, 2),
        "F1 Score": round(f1_score(y_true, y_pred, zero_division=0) * 100, 2)
    }
# Function
def run_fault_detection_analysis(dataframe, feature_cols, resolution_minutes=1):
    target_col = "Fault Detection Ground Truth"

    # 1.If we want to change the resolution we can
    if resolution_minutes > 1:
        dataframe = dataframe.iloc[::resolution_minutes, :]

    # 2.Variables
    X = dataframe[feature_cols].dropna()
    y = dataframe.loc[X.index, target_col]

    # 3. Scaler
    scaler = StandardScaler()
    X_scaled = scaler.fit_transform(X)

    # 4. Divide into train (70%) and test the other 30%
    X_train, X_test, y_train, y_test = train_test_split(X_scaled, y,
test_size=0.3, random_state=42)

    # 5. Models that are going to be used (DT, Random Forest and XGBoost )
    models = {
```

```
        "Decision Tree": DecisionTreeClassifier(random_state=42),
        "Random Forest": RandomForestClassifier(n_estimators=20,
random_state=42),
        "XGBoost": XGBClassifier(use_label_encoder=False, eval_metric='logloss',
verbosity=0, random_state=42)
    }

    # 6.Train the model
    results = {}
    for name, model in models.items():
        model.fit(X_train, y_train)
        y_pred = model.predict(X_test)
        results[name] = evaluate(y_test, y_pred)

    # 7. Resuslts showes in DataFrame
    results_df = pd.DataFrame(results).T
    return results_df
```

Run the fault detection analysis function (best feature for the first four PCs (Experimental))

```
features_pca = ['AHU: Return Air Fan Status', 'AHU: Outdoor Air Temperature',
'AHU: Return Air Temperature','AHU: Return Air Damper Control Signal']
results = run_fault_detection_analysis(dataexp, feature_cols=features_pca,
resolution_minutes=1)
results
```

SHAP Ranking:

```
##Experimental SHAP
target_col = "Fault Detection Ground Truth"
zone_cols = [col for col in dataexp.columns if col not in [target_col,
"Datetime"]]
# 1. Prepare features and target
X = dataexp[zone_cols]
y = dataexp.loc[X.index, target_col]

#  2. Scale and split
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)

X_train, X_test, y_train, y_test = train_test_split(X_scaled, y, test_size=0.3,
random_state=42)

#3.Train Models
#DT
dt_model = DecisionTreeClassifier(random_state=42)
dt_model.fit(X_train, y_train)
```

```
#Random Forest
rf_model = RandomForestClassifier(n_estimators=100, random_state=42)
rf_model.fit(X_train, y_train)

#XGBoost
xgb_model = XGBClassifier(use_label_encoder=False, eval_metric='logloss',
verbosity=0, random_state=42)
xgb_model.fit(X_train, y_train)

# 4. SHAP analysis using shap.Explainer()
explainer_rf = shap.Explainer(rf_model, X_train, feature_names=zone_cols)
shap_values_rf = explainer_rf(X_test,check_additivity=False)

explainer_xgb = shap.Explainer(xgb_model, X_train, feature_names=zone_cols)
shap_values_xgb = explainer_xgb(X_test,check_additivity=False)

explainer_dt = shap.Explainer(dt_model, X_train, feature_names=zone_cols)
shap_values_dt = explainer_dt(X_test,check_additivity=False)

#5. Summary plot example
shap.summary_plot(shap_values_rf, X_test, feature_names=zone_cols,
plot_type="bar", show=True)
shap.summary_plot(shap_values_xgb, X_test, feature_names=zone_cols,
plot_type="bar", show=True)
```

Code of F1 Score vs Number of Sensors (5 Best Features Experimental SHAP)

```
from scipy.optimize import curve_fit
results5['Number of Sensors'] = results5['Features'].apply(lambda s:
len(s.split(', ')))

# Group by # sensors and take the best F1 Score at each level
grouped = results5.groupby('Number of Sensors')['F1 Score'].max().reset_index()
x = grouped['Number of Sensors'].values
y = grouped['F1 Score'].values

# Define an exponential model and fit
def expo(x, a, b):
    return a * np.exp(b * x)

popt, _ = curve_fit(expo, x, y, p0=(y.min(), 0.01))
a, b = popt

# Compute R²
y_pred = expo(x, a, b)
ss_res = np.sum((y - y_pred)**2)
ss_tot = np.sum((y - np.mean(y))**2)
r2 = 1 - ss_res/ss_tot

# Plot
plt.figure(figsize=(8,6))
plt.scatter(x, y, s=50, label='Observed F1')
for xi, yi in zip(x, y):
```

```
    plt.text(xi, yi, f"{yi:.2f}%", verticalalignment='bottom',
horizontalalignment='center', fontsize=10)
plt.plot(x, y_pred, '--', label='Exponential fit')
plt.xticks(x)
plt.xlabel('Number of Sensors')
plt.ylabel('F1 Score (%)')
plt.title('F1 Score vs. Number of Sensors (5 Best Features Experimental SHAP)')
plt.grid(axis='y', linestyle='--', alpha=0.5)
plt.legend()

plt.tight_layout()
plt.show()
```