

**MÁSTER UNIVERSITARIO EN
Ingeniería de Telecomunicación**

TRABAJO FIN DE MÁSTER

***HERRAMIENTA INTERACTIVA PARA EL
DESARROLLO Y VISUALIZACIÓN DE ALGORITMOS
ADAPTADOS DEL SECTOR DE LA EFICIENCIA
ENERGÉTICA BASADOS EN TECNOLOGÍAS DE
MACHINE LEARNING Y ANALÍTICA DE DATOS***

**Alumna
Directora
Departamento**

*Lázaro, Elorriaga, Laura
Perea, Olabarria, Eugenio
Digital Energy, División de
Energía y Medio Ambiente,
Tecnalia.*

Curso académico

2017/2018

Bilbao, septiembre del 2018

Resumen

Debido a la preocupación actual por el ahorro energético y el efecto medioambiental de la generación de energía eléctrica, cada vez surgen mayor número de proyectos centrados en la generación de la energía a partir de fuentes renovables que buscan obtener mayor eficiencia en la producción y el consumo.

Este Trabajo Fin de Máster consiste en el diseño de una herramienta gráfica para el desarrollo y la visualización de algoritmos generados en la División de Energía y Medio Ambiente de la empresa Tecnalía y en el desarrollo de un nuevo algoritmo para el sector de la eficiencia energética. El proyecto se concreta en la aplicación de la herramienta y en la adaptación de los algoritmos para cuatro proyectos diferentes desarrollados por trabajadores de la división, los cuales ilustrarán diferentes aspectos de la eficiencia energética. Además, se desarrollará un algoritmo propio que hará uso de las tecnologías de *machine learning* y analítica de datos, tecnologías comúnmente utilizadas en el sector de la eficiencia energética, para desarrollar un algoritmo equivalente a uno de los algoritmos desarrollados en la empresa. El objetivo final de este trabajo fin de máster es que los trabajadores de la empresa sean capaces, haciendo uso de la herramienta diseñada, de exponer la principales características y ventajas de los diferentes proyectos que se han llevado a cabo en la empresa de una forma clara e interactiva.

Summary

Due to the current concern for energy saving and the environmental effect of electric power generation, a growing number of projects have been focusing on the generation of energy from renewable sources that seek greater efficiency in production and consumption.

The present Final Master Project consists on the design of a graphic tool for the development and visualization of adapted algorithms which have been carried out by workers of the Energy and Environment Division of the Tecnalia company and in the development of a new algorithm for the energy efficiency sector. The project is concreted in the design of the tool and in the adaptation of the algorithms for four different projects which have been developed by workers of the division and which will illustrate different aspects of energy efficiency. In addition, a new algorithm will be designed. It will make use of the most commonly used machine learning and data analytics technologies and it will be equivalent to one of the algorithms developed in the company. The final objective of this master's thesis is that the company's employees to be able, using the designed tool, to present the main characteristics and advantages of the different projects that have been carried out in the company in a clear and interactive way.

Laburpena

Energia aurrezteko eta energiaren sorkuntza elektrikoaren ingurumen-efektuari buruzko kezkek direla eta, energia berriztagarria ardatz duten gero eta proiektu gehiago garatzen hari dira, produkzioan eta kontsumoan eraginkortasun handiagoa bilatuz.

Master Amaierako Lan honen helburua, Tecnia enpresako Energia eta Ingurumen Dibisioan garatuak izan diren zenbait algoritmo garatzeko eta bistaratzeko gai den tresna grafiko bat garatzea eta eraginkortasun-energetikoari buruzko algoritmo berri bat egitea da. Proiektuak, Energia eta Ingurumen Dibisioan lan egiten duten langileak garatutako eraginkortasun energetikoaren alderdi ezberdinak islatzen duten lau proiektu desberdinentzako tresna grafikoaren aplikazioa garatzea eta algoritmoen egokitzapena lortzea du helburu. Horrez gain, *machine learning* eta datu analitikoaren teknologiaz baliatuta, enpresan garatutako algoritmoen baliokidea den algoritmo berri bat garatuko da. Helburu nagusia, enpresako langileek Master Amaierako Lan honetan diseinatutako tresna erabiliz, konpainian egin diren proiektuen ezaugarri eta abantaila nagusiak modu argi batean azaltzeko gai izatea da.

Contenido

Resumen.....	3
Summary	4
Laburpena.....	5
Contenido.....	6
Lista de Tablas	9
Lista de Ilustraciones	10
Lista de Acrónimos	12
1. Introducción	13
2. Contexto	15
2.1. Cambio Climático y Objetivo 20-20-20.....	16
2.1.1. Paquete de medidas sobre el clima y energía hasta 2020	16
2.1.2. Marco sobre clima y energía para 2030	17
2.2. Eficiencia Energética.....	17
2.3. Norma ISO 50001	17
2.4. Industria 4.0.....	19
2.4.1. Eficiencia Energética en la Edificación.....	20
3. Objetivos y Alcance	21
4. Beneficios	23
4.1. Beneficios Técnicos	23
4.2. Beneficios para Trabajadores y Clientes	23
4.2.1. Beneficios para la Empresa	23
4.2.2. Beneficios para los Clientes.....	23
4.2.3. Beneficios para el Alumno en Prácticas	24
4.3. Beneficios Económicos	24
5. Estado del Arte	25
5.1. Sistemas de Gestión	25
5.2.1. Sistemas de gestión de la Energía	25
5.2.2. Sistemas de gestión de Procesos.....	26
5.2. Tecnologías relacionadas	26

5.2.3.	Climatización	26
5.3.	Big Data	30
5.4.	Opciones de modelado.....	33
5.5.	Algoritmos Aprendizaje Máquina.....	34
5.5.1.	Aprendizaje Supervisado	36
5.5.2.	Aprendizaje no supervisado	38
6.	Análisis de Alternativas	40
6.1.	Obtención y Almacenamiento de los datos	40
6.2.	Programa en el que implementar la herramienta.....	41
6.3.	Proyectos que adaptar	42
6.4.	Algoritmos de analítica de datos.....	44
6.5.	Alternativas escogidas.....	46
7.	Descripción de las Tareas	47
7.1.	Recursos Utilizados.....	47
7.1.1.	Recursos Humanos	47
7.1.2.	Recursos Materiales	47
7.2.	Descripción de las Fases y Tareas. Gantt.....	48
	Fase 0: Gestión y Seguimiento del TFM	48
	Fase 1: Curso de Formación	48
	Fase 2: Documentación	49
	Fase 3: Estudio, Selección y Adaptación.....	49
	Fase 4: Diseño y Desarrollo de la Herramienta	50
	Fase 5: Simulaciones.....	51
	Fase 6: Redacción del Informe	51
	Duración total del trabajo	52
8.	Metodología	54
8.1.	Diseño de la herramienta	54
8.1.1.	Base de datos	54
8.1.2.	Demostradores.....	55
8.1.3.	Interfaz Gráfica	56
8.1.4.	Máquina Virtual.....	57
8.2.	Diseño del algoritmo	58
9.	Descripción de la Solución.....	61

9.1.	Demostrador #1: Zonas multitérmicas de un edificio terciario	61
9.2.	Demostrador #2. Previsión de carga de energía	65
9.3.	Demostrador #3. Despacho económico de la bomba de calor	67
9.4.	Demostrador #4. Sistema de diagnóstico de ineficiencias energéticas en industria.....	69
9.5.	Descripción del algoritmo desarrollado	71
9.5.1.	Librería SciPy	73
9.5.2.	Librería SkLearn	77
9.5.3.	Correlación	78
10.	Aspectos Económicos	81
10.1.	Horas internas	81
10.2.	Recursos materiales	82
10.2.1.	Amortizables.....	82
10.2.2.	Gastos.....	83
10.2.3.	Subcontrataciones.....	84
10.3.	Resumen.....	84
11.	Análisis de Riesgos.....	86
11.1.	Retrasos.....	86
11.2.	Bajas de participantes del proyecto	86
11.3.	Costes imprevistos.....	87
11.4.	Desarrollo de un Smart Energy Lab difícil de manejar	87
11.5.	Matriz probabilidad/ Impacto	88
12.	Conclusiones.....	89
13.	Bibliografía.....	91
14.	Anexos	94
	ANEXO I	95
	ANEXO II	98

Lista de Tablas

Tabla 1: Lista de procesos de análisis.....	35
Tabla 2: Peso de los criterios para la selección de la plataforma de trabajo.....	42
Tabla 3: Ponderación de la selección de la plataforma de trabajo.....	42
Tabla 4: Alternativas Escogidas para el desarrollo de la herramienta.....	46
Tabla 5: Recursos Humanos.....	47
Tabla 6: Recursos Materiales.....	47
Tabla 7: Fase1: Curso de formación.....	48
Tabla 8: Fase 2: Documentación.....	49
Tabla 9: Fase 3: Estudio, selección y adaptación.....	49
Tabla 10: Fase 4: Diseño y desarrollo de la herramienta.....	50
Tabla 11: Fase 5: Simulaciones.....	51
Tabla 12: Fase 6: Redacción del informe.....	51
Tabla 13: Horas Internas.....	81
Tabla 14: Fase de los recursos humanos para cada fase de trabajo.....	82
Tabla 15: Cálculo de la duración del TFG.....	82
Tabla 16: Cálculo de la vida útil de los materiales.....	83
Tabla 17: Cálculo de las amortizaciones.....	83
Tabla 18: Cálculo de los gatos.....	83
Tabla 19: Cálculo del presupuesto total.....	84

Lista de Ilustraciones

Ilustración 1: Logo Tecnalía (Fuente: Tecnalía)	15
Ilustración 2: Divisiones de negocio de Tecnalía (Fuente: Tecnalía)	15
Ilustración 3: Modelo de sistema de gestión de la energía para esta Norma Internacional (Fuente: ISO)	18
Ilustración 4: Tecnologías de la Industria 4.0 (Fuente: Smart Lighting)	19
Ilustración 5: Sistema HVAC residencial (Fuente: Gasfriocalor).....	27
Ilustración 6: Funcionamiento bomba de calor modo invierno (Fuente: Instalaciones Moreno & Gonzáles, S. L).....	29
Ilustración 7: Funcionamiento bomba de calor modo verano (Fuente: Instalaciones Moreno & Gonzáles, S. L).....	29
Ilustración 8: Diferentes modelos de Geotermia (Fuente: energiarenovablegeotermica).....	30
Ilustración 9: Big Data	31
Ilustración 10: Datos no Estructurados vs. Datos Estructurados (Fuente: Ndimensionz)	32
Ilustración 11: Principales Características Big Data (Fuente: 123RF).....	32
Ilustración 12: Modelo caja blanca (Fuente: virtual.itca.edu.sv)	33
Ilustración 13: Modelo caja negra (Fuente: virtual.itca.edu.sv)	34
Ilustración 14: Aprendizaje supervisado vs no supervisado. (Fuente: MathWorks).....	35
Ilustración 15: Aprendizaje supervisado. (Fuente: ResearchGate)	36
Ilustración 16: Algoritmo Random Forest. (Fuente: Youtube)	37
Ilustración 17: Algoritmo K-Nearest Neighbor (Fuente: Non-Parametric Techniques)	38
Ilustración 18: Aprendizaje no Supervisado. (Fuente: ResearchGate).....	38
Ilustración 19: Gant. Tiempo de dedicación previsto para las diferentes tareas y fases del TFM.....	53
Ilustración 20: Pasos seguidos para la realización de cada práctica	54
Ilustración 21: Datos del Museo Btek, cargados de la base de datos	56
Ilustración 22: Estructura Interfaz Gráfica Jupyter	57
Ilustración 23: Esquema Simplificado de una DMZ.....	58
Ilustración 24: Pasos seguidos para el diseño del algoritmo	58
Ilustración 25: Temperaturas para las 12 zonas térmicas.....	63
Ilustración 26: Agrupaciones realiza (con colores) de las diferentes zonas térmicas	63
Ilustración 27: Perfiles de Máximos y Mínimos de temperatura (%).....	64
Ilustración 28: Matriz de correlación. Similitud entre el comportamiento entre zonas térmicas.....	64
Ilustración 29: Diagrama de bloques del Escenario 2	65
Ilustración 30: Resultados para la simulación del escenario 2.....	66
Ilustración 31: Diagrama de bloques del escenario 3	67
Ilustración 32: Esquema del de Modelo de Control Predictivo.....	68
Ilustración 33: Resultados de simulación escenario 3.....	69
Ilustración 34: Procesamiento de los datos de entrenamiento	71
Ilustración 35: Mapa del edificio 700. Usado para la toma de los datos del Algoritmo I	71
Ilustración 36: Resultado para el Dendrograma y Método del codo	73
Ilustración 37: Agrupaciones con sistema de climatización encendido con algoritmo fcluster (maxclust).....	74

Ilustración 38: Resultados de la clusterización con fcluster de la librería SciPy para distancias de 0.5°C, 0.8°C, 1°C, 1.2°C y 1.8°C	75
Ilustración 39: Agrupación al aplicar fclusters con criterio de la distancia de 1°C	76
Ilustración 40: Clusterización de método k-Means de la librería SkLearn	77
Ilustración 41: Clusterización de métodos AgglomerativeClustering	78
Ilustración 42: Matrices de correlación de los tres periodos para método de fclusters con criterio MaxClust.....	79
Ilustración 43: Zonas con mayor y menos temperatura diaria (%)	80
Ilustración 44: Presupuesto final.....	85
Ilustración 45: Matriz Probabilidad-Impacto	88
Ilustración 46: Esquema de Bloques de la Base de datos	97
Ilustración 47: Menú de la Herramienta diseñada.....	99
Ilustración 48: Notebook del Escenario 1. Thermal Zone Analysis Module	100
Ilustración 49: Pasos para la ejecución del Notebook	101
Ilustración 50: Ejecución del Notebook. Introducción de parámetros para la simulación	102
Ilustración 51: Ejecución del Notebook. Visualización de los resultados.....	102
Ilustración 52: Reinicio del Kernel.....	103
Ilustración 53: Fin de la ejecución & Cierre de la aplicación.....	103

Lista de Acrónimos

TFM	Trabajo Fin de Máster
MV	Máquina Virtual
DMZ	Demilitarized Zone
URL	Uniform Resource Location
PC	Personal Computer
DEMA	División de Energía y Medio Ambiente
UE	Unión Europea
ISO	International Organization for Standarization
SGE	Sistemas de Gestión de la Energía
M&T	Monitorización y Telecontrol System
SCADA	Supervisory Control And Data Acquisition
PLC	Controlador Lógico Programable
HVAC	Heating, Ventilating, and Air Conditioning
TIC	Tecnologías de la Información y la Comunicación
TZAM	Modelo de Análisis de Zonas Térmicas
COP	Coficiente de Rendimiento
MCP	Modelo de Control de Predicción

1. Introducción

El presente Trabajo Fin de Máster (TFM) se ha desarrollado en el marco de un contrato de trabajo de la alumna Laura Lázaro con Tecnalía Research and Innovation, concretamente en la División de Energía y Medio Ambiente, en la Unidad Digital Energy.

Digital Energy es una Unidad que aglutina competencias transversales del mundo digital con aplicación al sector energético, como son la Inteligencia Artificial, Machine Learning y Analítica de Datos, la Realidad Virtual y Aumentada, la Ciberseguridad y el Blockchain. Se aplica a los subsectores energéticos como Operación y Mantenimiento de fuentes de generación renovable, Uso Final de la Energía y Eficiencia Energética, Gestión Óptima de las Redes de Distribución Eléctrica, Transformación de derivados del sector Oil and Gas (petroquímico) y Seguridad de Infraestructuras Energéticas.

Dentro de este marco de trabajo, Digital Energy está desarrollando un Digital Hub con demostraciones interactivas que permita a potenciales clientes del sector energético entender y palpar las ventajas de estas tecnologías digitales. En este TFM se ha desarrollado una herramienta para dicha Unidad, en concreto, dentro del Digital Hub, denominada Smart Energy Lab, que aglutina una serie de demostraciones para el subsector energético Uso Final de Energía de tecnologías Machine Learning y Analítica de Datos.

El Smart Energy Lab permitirá ejecutar diferentes demostraciones interactivas en las que se ejecutan algoritmos basados en técnicas de Analítica de Datos, diseñados e implantados en Digital Energy.

La alumna Laura Lázaro ha trabajado de forma individual, con el asesoramiento de los compañeros de Digital Energy, en la implementación de todos y cada uno de los algoritmos diseñados en Digital Energy en el Smart Energy Lab, y además ha aportado valor añadido en el diseño conceptual y desarrollo del algoritmo de zonas multitérmicas. La implementación del Smart Energy Lab incluye actividades como el desarrollo de la arquitectura del sistema y comunicaciones, desarrollo de la base de datos, desarrollo del entorno o framework de programas e interacción entre los mismos, comprensión y en un caso concreto, diseño de los algoritmos de analítica de datos, desarrollo de interfaces gráficas, análisis de la coherencia de resultados, etc... como se verá a lo largo del presente documento.

Para que compañeros de Tecnalía, a diferentes niveles, desde nivel técnico a directivo, puedan ejecutar las simulaciones del Smart Energy Lab en sus ordenadores de trabajo u ordenadores personales, la herramienta será instalada en una máquina virtual (MV) de la DMZ¹ de Tecnalía. De esta forma, una vez conectados a la red de la empresa, todo el mundo podrá ejecutarlo accediendo mediante una dirección web a la MV, sin necesidad de tener un programa de programación concreto instalado en su PC.

El propósito de este TFM ha sido el de diseñar una herramienta que haciendo uso de los Notebooks de Jupyter sirva como apoyo al personal de Tecnalía, de forma que se facilite la explicación y la comprensión de los diferentes proyectos que se llevan a cabo en la División de Energía y Medio Ambiente de la empresa. Será una herramienta utilizada por gestores de proyecto de Tecnalía para ilustrar de forma gráfica las ideas y conceptos principales de alguno de los proyectos desarrollados en la empresa y los diferentes algoritmos implantados en los mismos. Además, permitirá verificar el correcto funcionamiento de los algoritmos mediante la representación de los resultados en tiempo real.

¹ La DMZ o zona desmilitarizada es una zona insegura que se ubica entre la red interna de una organización y una red externa, generalmente en Internet.

2. Contexto

La Corporación Tecnalia [1] es una Corporación Tecnológica que nació en 2001 con el principal objetivo de contribuir al desarrollo del entorno económico y social a través del uso y fomento de la Innovación Tecnológica, mediante el desarrollo y la difusión de la investigación en un contexto internacional. La Corporación Tecnalia enmarca su actividad investigadora contribuyendo al fortalecimiento del desarrollo económico, la cohesión social y la sostenibilidad. Abierta a las colaboraciones, promueve relaciones y establece acuerdos con otros agentes de los sistemas de innovación.



Ilustración 1: Logo Tecnalia (Fuente: Tecnalia)

Tecnalia esta a su vez organizada en 7 divisiones de negocio. Entre ellas, la División de Energía y Medio Ambiente, DEMA, que se encarga de generar y desarrollar oportunidades de negocio de valor de los sectores de la energía y del medio ambiente, desarrollando productos, tecnologías y herramientas que puedan dar respuesta a diferentes necesidades.



Ilustración 2: Divisiones de negocio de Tecnalia (Fuente: Tecnalia)

2.1.Cambio Climático y Objetivo 20-20-20

Actualmente, la temperatura media mundiales es 0.85°C superior a la del siglo XIX [2]. Los estudios de clima realizados consideran que la actividad humana es la principal responsable de este incremento de temperatura. El ser humano está generando enormes cantidades de gases al quemar combustibles fósiles, talar las selvas tropicales y explotar ganado. Todo esto está repercutiendo en el clima y cambios de temperaturas que hacen que aumente el efecto invernadero y el calentamiento global.

Los gases de efecto invernadero son aquellos que retienen el calor del sol e impiden que se escape fuera. El CO₂ es un gas de efecto invernadero producido principalmente por la actividad humana y el responsable del 63% del calentamiento global causado por el hombre. Otros gases como el metano o el ácido nitroso, aunque se emitan en menores cantidades, son más potentes a la hora de retener el calor y por lo tanto más perjudiciales.

Debido a todo esto, la Unión Europea (UE) [3] se ha fijado objetivos para reducir progresivamente las emisiones de efecto invernadero de aquí a 2050. Para ello, la UE supervisa sus progresos en la reducción de las emisiones mediante actividades periódicas de seguimiento y elaboración de informes. Los principales objetivos climáticos y de energía quedan establecidos en:

2.1.1. Paquete de medidas sobre el clima y energía hasta 2020

Este paquete de medidas contiene legislación vinculante [4] que garantizará el cumplimiento de los objetivos climáticos y de energía asumidos por la UE para 2020. Las principales medidas son tres:

- 20% de reducción de las emisiones de gases de efecto invernadero respecto a los niveles de 1990.
- 20% de energías renovables en la UE.
- 20% de mejora de la eficiencia energética.

2.1.2. Marco sobre clima y energía para 2030

Los objetivos fundamentales del marco de clima y energía para 2030 son tres:

- Al menos 40% de reducción de las emisiones de gases de efecto invernadero respecto a los niveles de 1990.
- Al menos 27% de cuota de energías renovables.
- Al menos 27% de mejora de la eficiencia energética.

Entre las medidas que debe asegurar cada uno de los Estados, se incluyen estrategias para la movilización de inversiones con el objetivo de renovar los edificios existentes, tanto públicos como privados, la promoción de un uso eficiente de la energía por parte de los clientes finales, para ello los clientes recibirán contadores precisos a un precio competitivo, y el fomento de los servicios energéticos y su fácil acceso para las pequeñas y medianas empresas.

2.2. Eficiencia Energética

La Directiva de Eficiencia Energética [5] establece un conjunto de medidas vinculantes para ayudar a la UE a alcanzar su objetivo de eficiencia energética del 20% para 2020. Según la Directiva, todos los países de la UE deben utilizar la energía de manera más eficiente en todas las etapas de la cadena energética, desde la producción hasta el consumo final.

Como se ha mencionado anteriormente, entre las preocupaciones actuales está el ahorro de energía y el efecto medioambiental de la generación de energía eléctrica, buscando la generación a partir de energías renovables y una mayor eficiencia en la producción y el consumo, también denominado ahorro de energía. Por este motivo, la mayoría de los proyectos desarrollados por DEMA tiene como objetivo hacer un uso eficiente de la energía reduciendo así el consumo de la misma. De esta manera se busca optimizar los procesos productivos y el empleo de la energía utilizando lo mismo o menos para producir más bienes y servicios.

2.3. Norma ISO 50001

La norma ISO 50001 [6] es la norma internacional de sistemas de gestión de la energía (SGEn). Es una norma que incluye los requerimientos para definir, implementar, mantener y mejorar los sistemas de gestión de la energía con el propósito de permitir a una organización contar con un

enfoque sistemático para alcanzar una mejora continua en su desempeño energético, incluyendo la eficiencia energética, el uso y el consumo de la energía.

La Norma ISO 5001 se publicó oficialmente en 2011 con el objetivo de hacer un buen uso de la energía, para sí garantizar que las generaciones futuras puedan beneficiarse de ella y poder reducir las consecuencias del cambio climático. Surgió la necesidad de elaborar un estándar internacional que diera respuesta a las necesidades de las organizaciones y de la sociedad de un modo general.

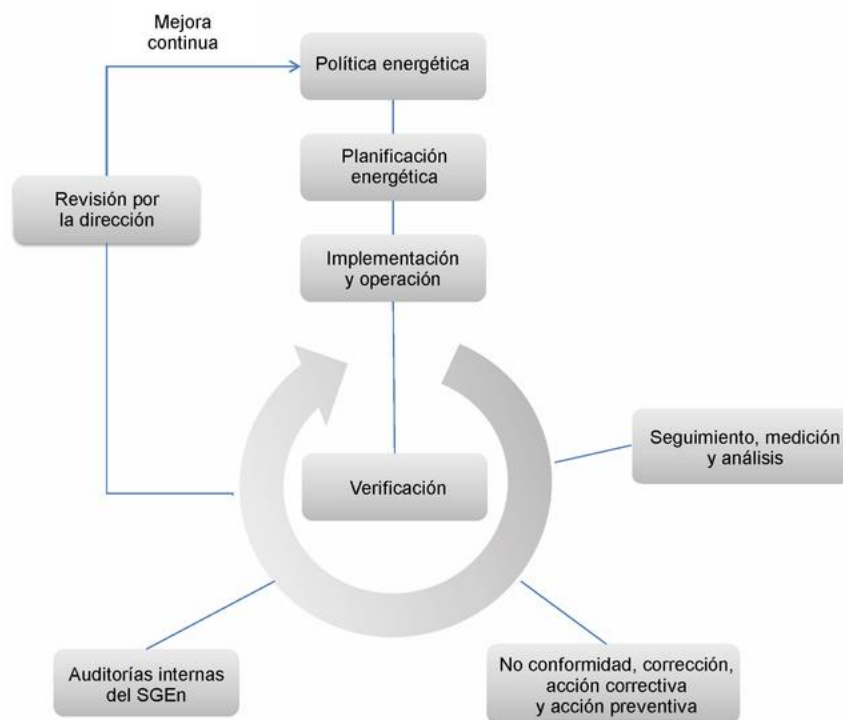


Ilustración 3: Modelo de sistema de gestión de la energía para esta Norma Internacional (Fuente: ISO)

Esta Norma Internacional especifica los requisitos aplicables al uso y consumo de la energía, incluyendo la medición, documentación e información, las prácticas para el diseño y adquisición de equipos, sistemas, procesos y personal que contribuyen al desempeño energético. Se basa en el ciclo de mejora continua: Planificar – Hacer– Verificar– Actuar (PHVA) e incorpora la gestión de la energía a las prácticas habituales de la organización tal como se ilustra en la Ilustración 3.

De cara a este TFM es particularmente relevante que la ISO 50001 impulsa la implantación de Sistema de Gestión Energética, favoreciendo la instalación de múltiples puntos de medida, y su almacenamiento. Registros de medida que el técnico de Eficiencia Energética debe analizar y

comprender. En este sentido las técnicas de Analítica de Datos son una herramienta tecnológica fundamental.

2.4. Industria 4.0

La Industria 4.0 [7], también llamada la Cuarta Revolución Industrial, es una nueva forma de organizar los medios de producción, mejorando la eficiencia energética de los equipos industriales. La asociación de las nuevas tecnologías de generación y eficiencia energética con las tecnologías inteligentes conduce a una industria con nuevas especializaciones productivas y centradas en la fabricación de productos eficientes que exige la reducción de costes energéticos y es más consciente del uso de la energía. La industria 4.0 se basa en el desarrollo de las siguientes 9 tecnologías:

- Big Data
- Cloud Computing
- Ciberseguridad
- Sistemas de integración horizontal y vertical
- Robótica
- Realidad aumentada
- 3D
- Simulación
- Internet of Things



Ilustración 4: Tecnologías de la Industria 4.0 (Fuente: Smart Lighting)

Para lograr una optimización de la eficiencia en la industria es muy importante lograr una buena monitorización de los consumos de los recursos en cada instalación y en su posterior análisis, de forma que se puedan determinar las oportunidades de mejora. Además, la conectividad también constituye un factor importante en la innovación en la mejora de la eficiencia. Mediante el Cloud Computing y la gestión de datos, se elevan las posibilidades del análisis de los datos recopilados.

2.4.1. Eficiencia Energética en la Edificación

La calefacción y la refrigeración de los edificios y de la industria representan la mitad del consumo de energía de la UE. Aproximadamente el 84% de la calefacción y refrigeración de los edificios se genera mediante combustibles fósiles, mientras que solo el 16% se genera a partir de energía renovable. Para cumplir los objetivos climáticos y energéticos de la UE, el sector de la calefacción y la refrigeración debe reducir el consumo energético y dejar de consumir combustibles fósiles. Debido a motivos como estos, las nuevas tecnologías asociadas a la Industria 4.0 se presentan, en el campo del ahorro y la eficiencia energética, como un factor clave en el cambio de modelo productivo.

3. Objetivos y Alcance

El presente Trabajo Fin de Máster tiene dos objetivos principales:

1. Por un lado, el diseño e implementación de un Smart Energy Lab para Digital Energy de Tecnalia dentro de su Digital Hub que, haciendo uso de la aplicación web Jupyter, sirva para ilustrar los diferentes proyectos llevados a cabo por la Unidad. Proyectos todos ellos que se basan en manipular datos de un set de medidas importante (Big Data) registrados en una Base de Datos, para extraer conocimiento y comprender los patrones de consumo bien para la caracterización de los mismos hacia la Eficiencia Energética (descriptivo), hacia la predicción de consumos (predictivo), o la sugerencia de consignas de operación de las infraestructuras energéticas (prescriptivo).
2. Y por el otro, el desarrollo de un algoritmo para la caracterización de zonas multitérmicas en el edificio (descriptivo) que, haciendo uso de técnicas de analítica de datos, caracterice y agrupe en base a la variación con la temperatura de las diferentes zonas térmicas del edificio.

Con el fin de que sirva como apoyo en las explicaciones a posibles clientes de la empresa, se van a adaptar los diferentes algoritmos que cubran en una amplia medida la mayoría de los conceptos trabajados en Digital Energy de Tecnalia y se van a diseñar diferentes herramientas gráficas para ilustrar los resultados. La herramienta permitirá la visualización de los resultados obtenidos y calculará, mediante algoritmos de analítica de datos, soluciones de ahorro energético.

Fundamentalmente será una herramienta utilizada por gestores de proyectos de la División de Energía y Medio Ambiente de Tecnalia para ilustrar de forma gráfica las ideas en las que se basan los proyectos que se presenten. También se dispondrá de métricas para evaluar los resultados obtenidos.

Además de los objetivos principales, existen otros objetivos secundarios detallados a continuación, que ayudarán a que el objetivo principal se cumpla satisfactoriamente:

- Dimensionado para 4 demostraciones de desarrollos Smart Energy Lab. Después de estudiar diferentes alternativas, se ha decidido que los proyectos que se desarrollarán serán los siguientes:
 1. Zonas multitérmicas de un edificio terciario
 2. Previsión de carga de energía.
 3. Despacho económico de la bomba de calor.
 4. Sistema de diagnóstico de ineficiencias energéticas en industria
- Adaptar todos los algoritmos para que todos estén accesibles en lenguaje Python.
- Adaptar todos los algoritmos para que todos estén diseñados en base a una estructura similar.
- Desarrollar una herramienta visual e interactiva, con el objetivo de facilitar el manejo a los usuarios.
- Desarrollar una herramienta que haga un cálculo sistemático y una presentación amigable de resultados, haciendo visibles los resultados de una manera rápida y fácil de entender.
- Desarrollar una herramienta que, al estar implantada en un Máquina Virtual de la DMZ de Tecnalia, sea accesible por cualquier trabajador de la empresa mediante una URL dada y sin necesidad de tener ningún programa especial instalado en su PC.

En este Trabajo Fin de Máster, además de explicar el algoritmo desarrollados por el alumno, se entregará una breve contextualización de cada uno de los proyectos que se han desarrollado junto con su respectiva interfaz gráfica y una breve demostración del correcto funcionamiento de cada una de ellas. Con el fin de facilitar el uso de la herramienta, también se entregará anexo a la memoria un manual de usuario.

4. Beneficios

4.1. Beneficios Técnicos

La realización de presente Trabajo Fin de Máster dará lugar a una serie de beneficios técnicos como son:

- Creación de un centro de demostración de competencias.
- Creación de un entorno de trabajo para la simulación experimental.
- Consolidación de datos de consumo y proceso de diferentes sistemas energéticos en una base de datos centralizada.
- Adecuación de los códigos de los diferentes proyectos para que todos ellos estén desarrollados en base a una misma estructura y así tengan un esqueleto similar.

4.2. Beneficios para Trabajadores y Clientes

4.2.1. Beneficios para la Empresa

La herramienta en la que se ha trabajado servirá como apoyo a los trabajadores de Tecnalía. Haciendo uso de esta herramienta se pretende facilitar y completar las explicaciones sobre los diferentes proyectos en los que se está trabajando.

El hecho de trabajar con una herramienta que posibilite la simulación de los diferentes escenarios dará la oportunidad de aplicar diferentes algoritmos de procesado y realizar la evaluación posterior de los resultados obtenidos. Además, gracias a las diferentes interfaces interactivas que se han desarrollado, se podrán variar los parámetros y ver cómo afectan a los resultados de una forma rápida y sistemática. Estos resultados podrán ser almacenados para una futura visualización.

4.2.2. Beneficios para los Clientes

La herramienta que se va a realizar será también beneficiosa para los clientes ya que les ayudará a entender mejor los desarrollos presentados por la empresa y obtendrán un resultado visual y a tiempo real de los mismos. Servirá para que el potencial cliente

tenga una fácil comprensión del alcance y posibilidades de las nuevas tecnologías digitales, más concretamente, de la Inteligencia Artificial en su negocio. Además, mediante las simulaciones quedará verificado el correcto funcionamiento de los algoritmos realizados, ya que los datos serán obtenidos de la base de datos que recogerá los valores de los diferentes sensores de los edificios.

4.2.3. Beneficios para el Alumno en Prácticas

El trabajar con diferentes algoritmos implantados en diferentes escenarios permitirá al alumno aprender acerca de nuevos métodos de eficiencia y gestión de la energía y aprender más acerca de algoritmos de aprendizaje máquina.

Además, el hecho de que los algoritmos estén desarrollados en diferentes lenguajes de programación conllevará a que el alumno se familiarice con nuevos lenguajes como Python, C++ o Java.

4.3. Beneficios Económicos

La herramienta implicará por tanto beneficios económicos. Mediante la implantación del Smart Energy Lab en uno de los laboratorios de simulación de los que dispone la empresa en el edificio 700 de Derio, Bizkaia, se podrán guardar los datos recogidos por los sensores que estén disponibles en el laboratorio en la base de datos para así luego poder analizar y visualizar los resultados a tiempo real. El hecho de probar el correcto funcionamiento de los diferentes algoritmos aumentará el grado de confianza de los potenciales clientes hacia la empresa.

Además, se espera que el hecho de usar la herramienta para completar las explicaciones, modificar parámetros de las simulaciones e ilustrar gráficamente los resultados obtenidos con los diferentes algoritmos haga que los potenciales clientes se sientan más atraídos por los diferentes proyectos desarrollados.

5. Estado del Arte

Para poder desarrollar los algoritmos de un modo eficaz, y haciendo uso de las tecnologías disponibles hoy en día, ha sido necesario hacer un estudio tanto de los diferentes mecanismos de gestión de energía disponibles, como de los diferentes mecanismos de algoritmia disponibles para procesar los datos obtenidos de la base de datos.

5.1. Sistemas de Gestión

5.2.1. Sistemas de gestión de la Energía

La Norma UNE-ISO 50001 [8] define los Sistemas de Gestión de la Energía como el conjunto de elementos interrelacionados o que interactúan para establecer una política y objetivos energéticos, y los procesos y procedimientos necesarios para alcanzar dichos objetivos.

Los Sistemas de Gestión Energética (SGEn) [9] responden a las necesidades propias de un Gestor Energético, facilitando la información previamente tratada y presentada de forma que permiten la extracción inmediata de conclusiones.

La necesidad de las empresas de conocer los costes reales de la energía que consumen, sumado con la necesidad de realizar auditorías de eficiencia energética con regularidad, lo cual viene dictaminado por la Directiva de Eficiencia Energética de la Unión Europea, han hecho que los sistemas de gestión de la energía sean cada vez más necesarios. Las empresas podrán reducir su consumo energético para disminuir los costes operativos y así promover la sostenibilidad económica, política y ambiental. Además, podrán aumentar su eficiencia para maximizar su beneficio mientras que reducen las emisiones de gases de efecto invernadero.

Los sistemas de supervisión y monitorización hacen posible mantener un rendimiento constante a la vez que aseguran una correcta gestión energética. Estos sistemas capturan y ofrecen datos para identificar oportunidades de ahorro y detectar ineficiencias en los procesos optimizando así las facturas energéticas. Por lo general,

estos sistemas están compuestos por dos partes, Monitorización y Telecontrol (M&T) y SGEEn.

Los sistemas de Monitorización y Telecontrol ofrecen información de diferentes magnitudes, como son los parámetros de consumo, para su posterior tratamiento y análisis. Suelen estar basados en sistemas SCADA, de forma que los Sistemas de Gestión Energética puedan apoyarse en sus datos. El termino SCADA, Supervisor y Control And Data Acquisition, hace referencia al software para ordenadores que permite controlar y supervisar sistemas industriales a distancia.

5.2.2. Sistemas de gestión de Procesos

Los sistemas de gestión de procesos [10] son sistemas basados en dispositivos diseñados para centralizar la recolección de los datos de los procesos en una planta, ejecutar programas de control y realizar acciones sobre los procesos en forma autónoma a través de interconexiones con la instrumentación y con otros dispositivos de supervisión y control. Los más comunes son los Controladores Lógicos programables (PLC, por sus siglas en inglés). Los Controladores Lógicos programables son dispositivos de estado sólido, basado en Microprocesadores, utilizados para controlar la operación de una máquina, proceso o planta por intermedio de un programa o algoritmo almacenado, recibiendo información realimentada desde el proceso mediante instrumentos y dispositivos de entrada y salida.

5.2. Tecnologías relacionadas

5.2.3. Climatización

La climatización comprende tres cuestiones fundamentales: la ventilación, la calefacción., o climatización de invierno, y la refrigeración climatización de verano. Los sistemas de climatización se basan en el tratamiento de aire para obtener un control simultáneo de la humedad, la temperatura y de una correcta distribución en los espacios interiores como una habitación o en un edificio.

Aunque existen varios tipos de sistemas de climatización, a continuación, se va a explicar el sistema Heating, Ventilation and Air Conditioning (HVAC) [11], por ser el más completo y el más común en los edificios en los que se va a trabajar.

HVAC

Heating, Ventilation and Air Conditioning, es una tecnología de confort ambiental interior y exterior que tiene como objetivo proporcionar confort térmico y una calidad de aire interior aceptable. El diseño del sistema HVAC es una disciplina de ingeniería mecánica, basada en los principios de termodinámica, mecánica de fluidos y transferencia de calor. En la siguiente imagen se muestra un sistema HVAC residencial:

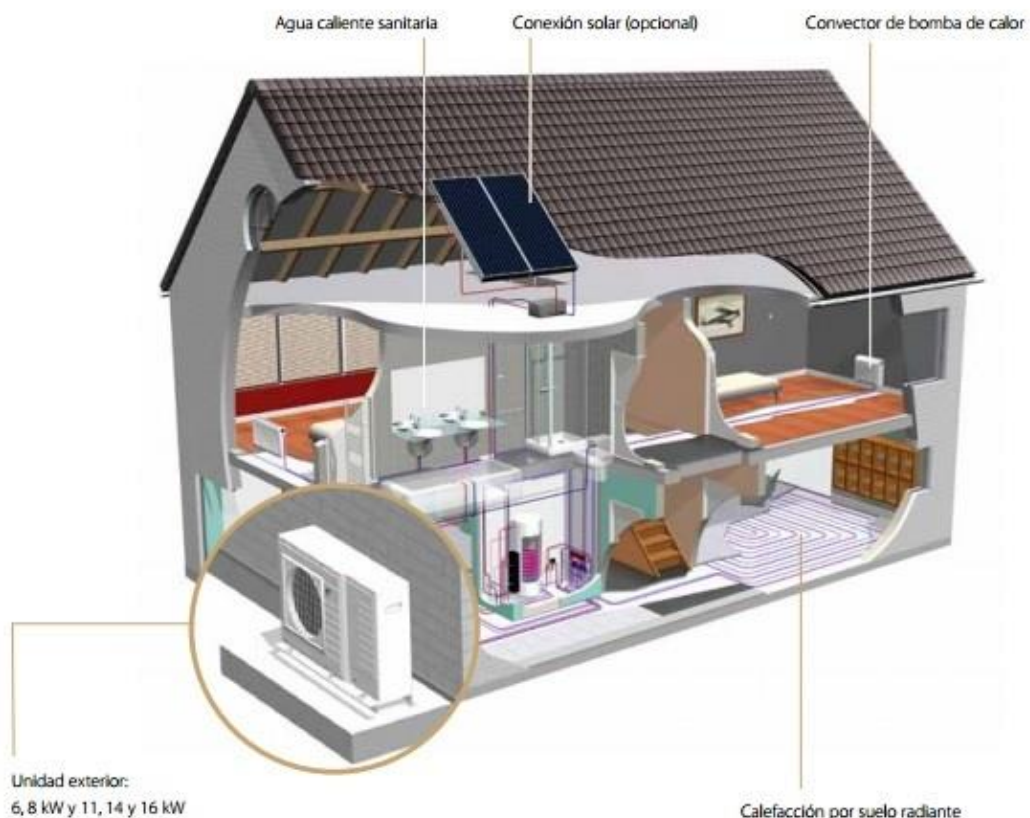


Ilustración 5: Sistema HVAC residencial (Fuente: Gasfriocalor)

Las bombas de calor son la principal tecnología utilizada para la generación de la energía térmica, ya que pueden funcionar a modo de calefacción como a modo de refrigeración, en función de la temperatura del fluido que se le introduzca.

Bombas de Calor

Las bombas de calor [12] son un sistema de calefacción que absorbe el calor presente en un fluido exterior y lo transporta hacia un ambiente interior a climatizar. La principal fuente de energía proviene del calor existente en el aire o en la tierra, por ello se trata de un sistema de bajo consumo y ecológico. Las bombas de calor son equipos reversibles, es decir, pueden refrigerar o calentar el edificio, dependiendo de las condiciones exteriores.

Los diferentes ciclos de la bomba de calor son los siguientes:

1. Compresión: El compresor transforma en calor la energía eléctrica y lo transmite al fluido refrigerante para elevar su presión y su temperatura.
2. Condensación: El refrigerante pasa de estado gaseoso a estado líquido mediante el intercambiador de calor, donde el fluido cede su calor al foco caliente y al enfriarse se produce la reacción exotérmica.
3. Expansión: El fluido pasa por un dispositivo que genera una pérdida de carga por lo que se produce una reducción de presión, reduciéndose la temperatura.
4. Evaporación: El evaporador está situado en la fuente fría y es donde se produce una reacción endotérmica (de absorción de calor) y se vuelve a iniciar el primero de los ciclos.

En las siguientes dos ilustraciones se muestran los diferentes ciclos de la bomba de calor. En la segunda imagen se puede apreciar cómo, el funcionamiento de la bomba de calor es inverso cuando se trabaja para generar frío:

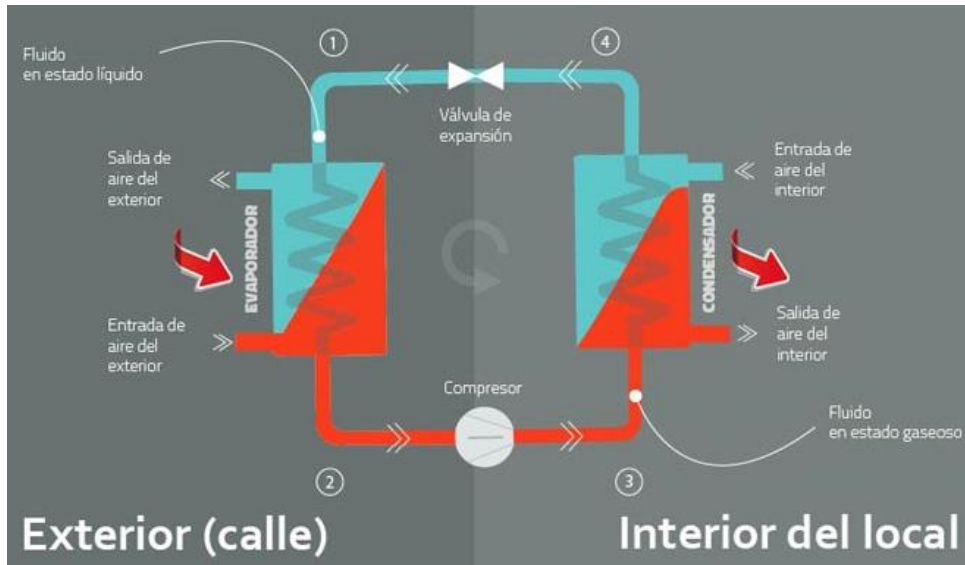


Ilustración 6: Funcionamiento bomba de calor modo invierno (Fuente: Instalaciones Moreno & Gonzáles, S. L)

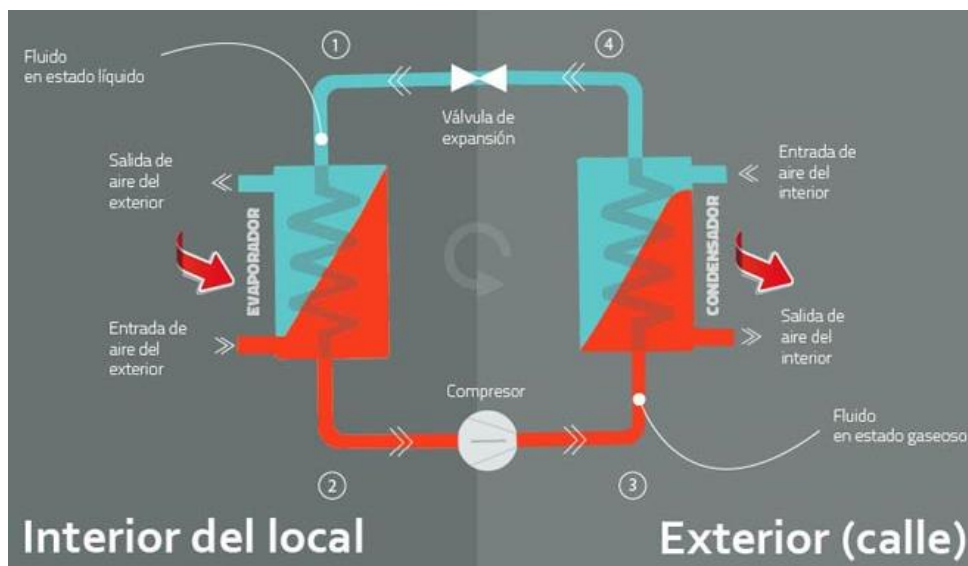


Ilustración 7: Funcionamiento bomba de calor modo verano (Fuente: Instalaciones Moreno & Gonzáles, S. L)

En cuanto a la clasificación, las bombas de calor pueden clasificarse en tres grandes grupos según el origen y el medio de la energía: aerotérmicas, geotérmicas e hidrónicas[13].

Aerotérmica

- Bombas de calor aire-aire: Transforman el aire de una temperatura en otra.
- Bombas de calor agua-aire: Aprovechan la energía que contiene el agua de los ríos y mares.

Hidráulicas

- Bombas de calor aire-agua: extraen la energía del aire exterior y lo convierten en frío, calor y agua sanitaria.
- Bombas de calor agua-agua: Extraen el calor para la calefacción del agua subterráneo con radiadores de baja temperatura, venticonectores o suelos radiantes.

Geotérmicas

- Bombas de calor tierra-aire: Utilizan las propiedades de la tierra para aportar calor y frío al edificio.
- Bombas de calor tierra-agua: Aprovechan el calor contenido en el terreno y mediante una condensación por agua aportan calor o frío al edificio.

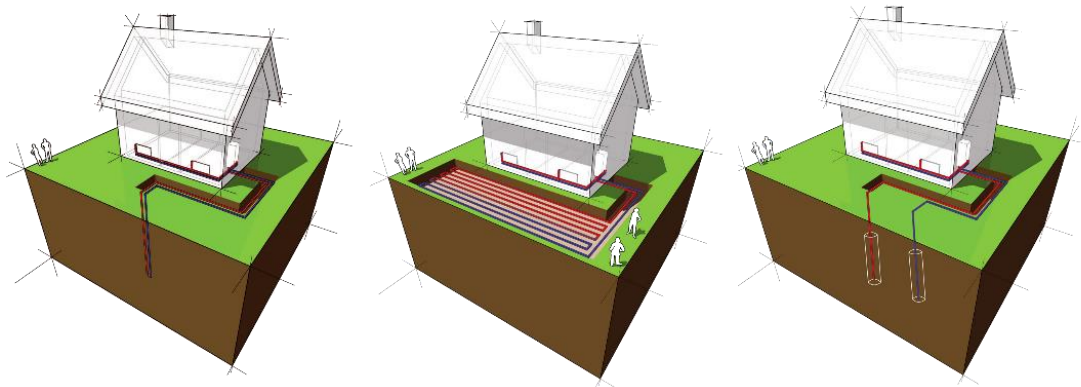


Ilustración 8: Diferentes modelos de Geotermia (Fuente: energiarenovablegeotermica)

5.3. Big Data

Big Data [14] o datos a gran escala hace referencia a un conjunto de datos tan grande que las aplicaciones informáticas tradicionales de procesamiento de datos no son capaces de tratar con ellos ni de encontrar patrones repetitivos. Se encuentra dentro del sector de las tecnologías de la información y la comunicación (TIC) y se ocupa de la manipulación y procesamiento de grandes volúmenes de datos.

Nació con el objetivo de cubrir la problemática existente del almacenamiento, tratamiento y aprovechamiento de los grandes volúmenes de datos que se producen en la actualidad por factores como son: la elevada y creciente cantidad de fuentes de datos (sensores y redes sociales, por ejemplo) y la generalización de las redes de telecomunicaciones, en muchos casos inalámbricas. El conjunto de estos elementos, junto con las mayores capacidades de almacenamiento, ha hecho crecer de una manera enorme la cantidad de datos disponibles en los últimos años, tendencia que se sigue manteniendo en la actualidad.



Ilustración 9: Big Data

Dichos volúmenes de datos poseen cuatro características principales que vienen definidas como las cuatro Vs [15]:

- *Volumen de información*
Cantidad de datos que son generados a lo largo del tiempo. Es una de las principales características de Big Data, ya que hace referencia a las cantidades masivas de datos que se almacenan para ser procesados.
- *Velocidad de los datos*
Rapidez con la que los datos son creados, almacenados y procesados en tiempo real. En muchas ocasiones es necesario hacer un estudio en tiempo real.
- *Variedad de los datos*
Formas, tipos y fuentes en las que los datos son registrados. Los datos pueden ser estructurados y fáciles de gestionar como son las bases de datos, o no estructurados, como son los documentos de texto, correos electrónicos, datos de sensores, etc.

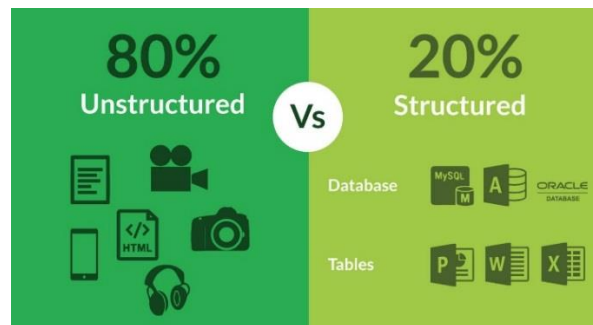


Ilustración 10: Datos no Estructurados vs. Datos Estructurados (Fuente: Ndimensionz)

Los datos no estructurados requieren un tratamiento diferente de los datos no estructurados, ya que es necesario un procesamiento de los datos recogidos de múltiples fuentes de información con la herramienta adecuada.

- Veracidad de los datos

Grado de falibilidad de los datos recibidos. Es necesario tener la certeza de que los datos obtenidos son de calidad, aplicando soluciones y métodos que puedan eliminar datos imprevisibles.



Ilustración 11: Principales Características Big Data (Fuente: 123RF)

A estas cuatro V, se le pueden añadir otras tres más:

- Viabilidad

La capacidad de generar un uso eficaz de los volúmenes de datos. Esta característica también se asocia a la innovación de los equipos de trabajo y el uso de tecnologías empleadas. Tras haber analizado, seleccionado y monitorizado la información, es necesario ser capaz de filtrar y seleccionar los atributos y factores de interés para el caso de estudio.

- Visualización de los datos

Una vez que los datos han sido procesados, es necesario tener un modo en el que poder representarlos, de forma de que sean legibles y accesibles por los usuarios. Existen herramientas de visualización que ayudan a comprender los datos gráficamente.

- Valor de los datos

El valor de los datos que se transformación en información.

5.4. Opciones de modelado

Dependiendo del a cantidad de información y de los datos disponibles, los problemas de modelado se clasifican en tres grupos [16]:

Caja Blanca

Se denomina modelado de caja blanca cuando tanto la arquitectura como los datos del sistema son conocidos. Se tienen en cuenta las entradas y las salidas, así como la interacciones, las conexiones interiores y las relaciones entre los posibles subsistemas. Este tipo de modelado no facilita el testeado entre diferentes módulos. Debido a los usos impredecibles de los edificios, este tipo de modelado es menos preciso que el resto.

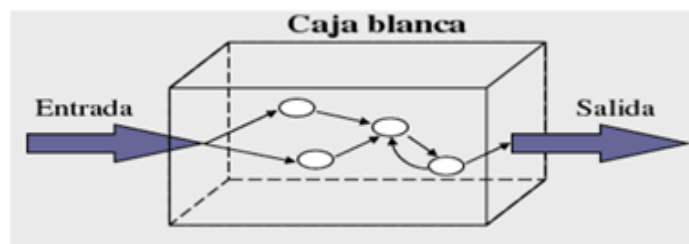


Ilustración 12: Modelo caja blanca (Fuente: virtual.itca.edu.sv)

Caja Gris

Sistemas de los que se tiene algún tipo de conocimiento de su sistema interno.

Caja Negra

El modelo de caja negra se basa en la idea de probar un sistema sin tener ningún conocimiento del funcionamiento interno del sistema o de su arquitectura. Solo interesan, por tanto, las entradas y salidas de materia, la energía e información en el sistema, y no los elementos e interacciones que suceden en el interior. Estos tipos de modelado son bastante útiles en la configuración de reconocimiento de patrones. A diferencia de las pruebas de caja blanca, las pruebas de caja negra facilitan la comunicación de prueba entre los módulos.

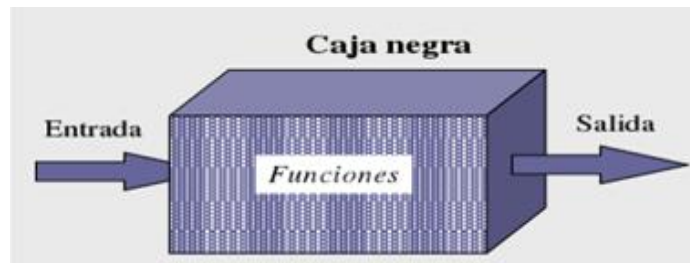


Ilustración 13: Modelo caja negra (Fuente: virtual.itca.edu.sv)

Haciendo uso de diferenciadas técnicas de aprendizaje máquina algunos algoritmos se pueden considerar como una caja negra que son capaces de predecir un patrón de salida cuando reconoce un patrón de entrada dado.

5.5. Algoritmos Aprendizaje Máquina

Los problemas de Aprendizaje Máquina tratan de predecir las propiedades o el modo de trabajo de una colección de datos seleccionados [17]. Los sistemas de aprendizaje pueden ser separados en dos grupos: aprendizaje supervisado y aprendizaje no supervisado [18].

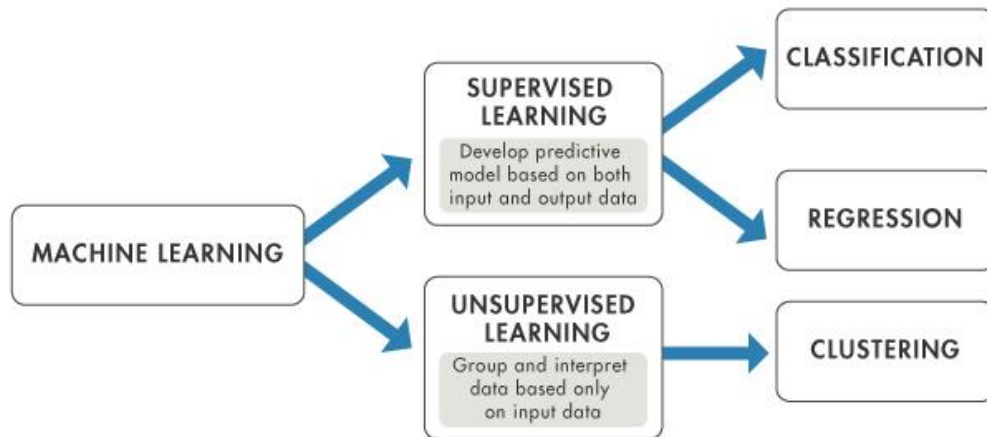


Ilustración 14: Aprendizaje supervisado vs no supervisado. (Fuente: MathWorks)

La siguiente tabla presenta una lista de procesos de análisis por tipo. A continuación, se explicarán los que alguna vez se consideraron de mayor importancia y más relevantes para el trabajo que se ha realizado:

ALGORITMO	CATEGORIA
METODOS DE CLUSTERING	
(1) K-Means	Clustering
(2) Gaussian Mixtures	Clustering
METODOS DE CLASIFICACIÓN/ REGRESIÓN	
(3) Ordinary Least Squares (OLS)	Regresión, Selección de Características
(4) Naive Bayes (NB)	Clasificación
(5) K-Nearest Neighbours (k-NN)	Clasificación, Regresión
(6) Support vector machines (SVM)	Clasificación, Regresión
(7) Decision Trees (DT)	Clasificación, Regresión
(8) Random Forest (RF)	Clasificación, Regresión
(9) Recurrent Neural Networks (RNN)	Clasificación, Regresión

Tabla 1: Lista de procesos de análisis

5.5.1. Aprendizaje Supervisado

En el aprendizaje supervisado, la función se genera a partir de datos de entrenamiento o de un conjunto de ejemplos entrenados. El algoritmo analiza los datos de entrenamiento y produce una función que se usará para mapear datos nuevos. Es necesario trabajar con un algoritmo de aprendizaje que permita llegar, partiendo de los datos de entrenamiento, a posibles situaciones futuras. Este problema puede ser resuelto por:

- Métodos de clasificación:** consiste en la identificación de la categoría a la que pertenece una nueva observación de un conjunto de datos de capacitación. En otras palabras, comparar los datos ya etiquetados con los nuevos datos no etiquetados para predecir a qué categoría pertenecen.
- Métodos de regresión:** consiste en la estimación de la relación entre variables. Las técnicas de regresión ayudan a comprender cómo cambia el valor de una variable dependiente cuando varía otra variable independiente.

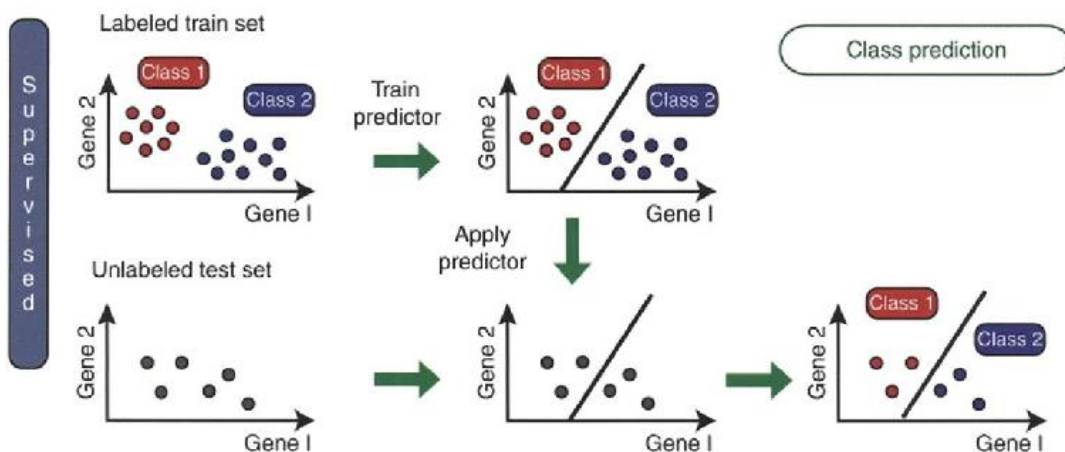


Ilustración 15: Aprendizaje supervisado. (Fuente: ResearchGate)

Algoritmo de árbol de decisión

Los árboles de decisión [19] son un tipo de aprendizaje supervisado donde un conjunto de condiciones se organiza jerárquicamente. El objetivo es crear un modelo que prediga el valor de una variable mediante el aprendizaje de reglas simples de decisión inferidas a partir de las características de los datos. De esta manera, es posible ir desde la raíz del árbol hasta las hojas para llegar a la decisión final.

Los modelos de árbol donde el objetivo puede tomar un conjunto discreto de valores se denominan árboles de clasificación. Los árboles de decisión donde el objetivo puede tomar valores continuos se llaman árboles de regresión.

- **Random Forest**

Random Forest es una combinación de predictores de árbol. Cada árbol depende de los valores de una muestra de vector aleatorio de forma independiente y con la misma distribución que el resto de los árboles del bosque. El Random Forest comienza con una técnica de aprendizaje automático estándar llamada "árbol de decisiones", que, corresponde a un aprendizaje. En un árbol de decisión, una entrada se introduce en la parte superior y hacia abajo a medida que atraviesa el árbol de los datos se acumulan en conjuntos cada vez más pequeños. En la siguiente ilustración se muestra el principal funcionamiento del algoritmo.

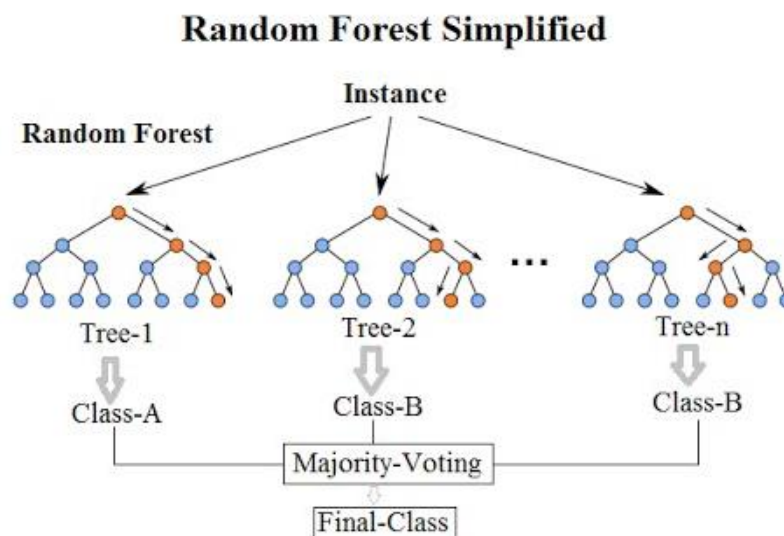


Ilustración 16: Algoritmo Random Forest. (Fuente: Youtube)

[Algoritmo K-Nearest Neighbor](#)

El algoritmo de vecino más cercano es un método utilizado para la clasificación y la regresión de reconocimiento de patrones. En la clasificación K-NN, las nuevas muestras se clasifican calculando la distancia al caso de entrenamiento más cercano y luego, el objeto se clasifica en función del voto mayoritario de sus vecinos. En la

regresión K-NN, el valor de la salida es el promedio de los valores de los K vecinos más cercanos.

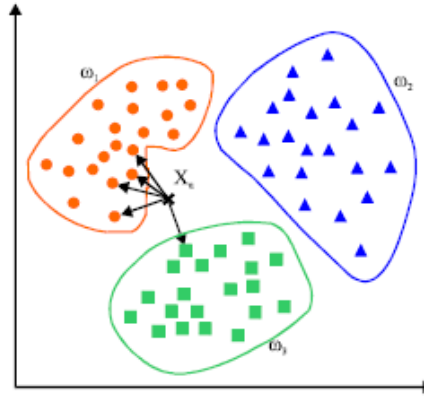


Ilustración 17: Algoritmo K-Nearest Neighbor (Fuente: Non-Parametric Techniques)

5.5.2. Aprendizaje no supervisado

En el aprendizaje no supervisado, no existe un conjunto de datos de entrenamiento. Consisten en un conjunto de vectores de entrada sin ningún valor anotado y los resultados son desconocidos. Funciona solo con datos de entrada y operaciones lógicas para resolver problemas complejos; sin ningún dato de referencia. Una forma típica de resolver los problemas es mediante la agrupación o clustering.

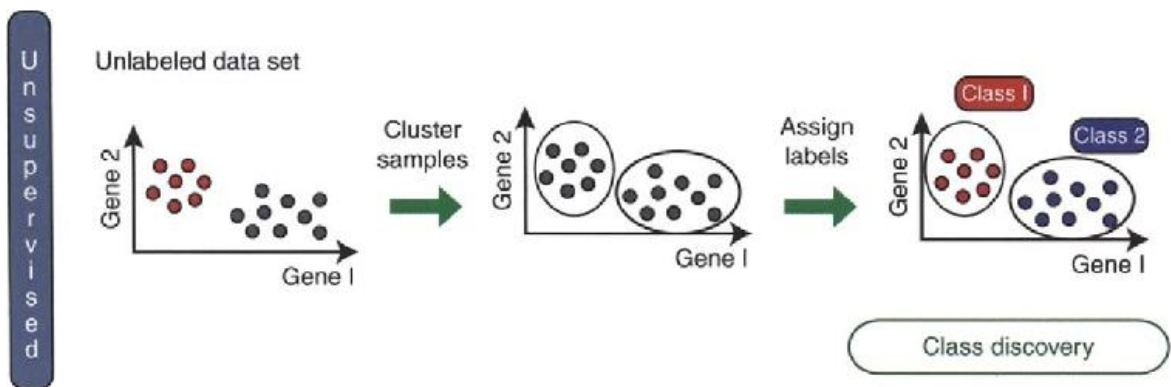


Ilustración 18: Aprendizaje no Supervisado. (Fuente: ResearchGate)

- Métodos de agrupación:** consiste en agrupar objetos similares en un mismo grupo creando diferentes sets. Dado un conjunto de puntos de datos, se crean un conjunto de grupos donde cada punto dentro de cada grupo sea similar entre sí.

Red neuronal artificial

Durante el entrenamiento de una red neuronal artificial, bajo el aprendizaje no supervisado, los vectores de entrada de tipo similar se combinan para formar clúster. Cuando se aplica un nuevo patrón de entrada, la red neuronal da una respuesta de salida que indica la clase a la que pertenece el patrón de entrada.

En las redes neuronales no supervisadas, no hay ningún indicador para saber si el resultado es correcto o no. Por lo tanto, es la red en sí la responsable de descubrir los patrones, presentando datos de entrada y presentando la relación de los datos de entrada sobre la salida.

Algoritmo k-Means

K-means es un algoritmo de clasificación no supervisada (clusterización) que agrupa objetos en k grupos basándose en sus características. El agrupamiento se realiza minimizando la suma de distancias, normalmente cuadrática, entre cada objeto y el centroide de su grupo o clúster.

6. Análisis de Alternativas

Para el desarrollo de este trabajo, ha procedido al estudio y selección de varias alternativas de diferentes aspectos. En este apartado se analizan dichas alternativas y se explican los motivos de su elección.

6.1. Obtención y Almacenamiento de los datos

En cuanto a la adquisición de los datos, cada uno de los diferentes escenarios requerirá una fuente diferente de recopilación de datos. Cada demostrador hace uso de sensores de diferentes fabricantes para la obtención de las medidas de consumo energético o de proceso (bomba de calor, climatización en edificio, o conformado de piezas en industria).

Se han estudiado los diferentes proyectos con los que se va a trabajar y se ha decidido cuál será la mejor fuente de donde adquirir los datos. Debido a la importancia del correcto almacenamiento de los datos, y para que en un futuro sea posible disponer de ellos de forma sencilla, se han estudiado dos alternativas:

- *Almacenamiento en BD diferentes*

El hecho de tener los datos en BD diferentes posibilita la opción de que cada base de datos pueda estar estructura de forma diferente. De esta forma, cada edificio o proyecto podría disponer una base de datos particular. Por otro lado, cada algoritmo quedará restringido de datos en los que poder ser probados.

- *Almacenamiento en una única BD*

Tener todos los datos en una misma base de datos supondrá que todos los datos estén estructurados de forma similar. Además, dará la opción de poder probar los diferentes algoritmos de diferentes proyectos con diferentes grupos de datos, para poder así testear el funcionamiento de los algoritmos desarrollados en un modo más amplio.

Aunque en un principio sea la opción que más trabajo requiera, por estar los datos de diferentes proyectos en diferentes BD, se ha decidido agrupar todos los datos en una única base de datos (de nombre efiDB), de forma que todos los datos puedan ser usados para

todos los escenarios desarrollados para la herramienta. Los datos de los diferentes sensores requeridos para los diferentes algoritmos se irán volcando a esta base de datos de forma periódica.

6.2. Programa en el que implementar la herramienta

Una vez fijado cual es el propósito y el alcance de la herramienta que se desea implementar, se ha estudiado diferentes programas de programación para ver cuál es el que mejor se ajusta a los objetivos fijados anteriormente.

Se han estudiado dos opciones posibles:

- 1 Aplicación web Jupyter [20],
- 2 Software con licencia Matlab [21].

Los criterios de selección de estas dos alternativas son:

- Licencia. La aplicación web Jupyter existe para desarrollar software de fuente abierta, estándares abiertos y servicios para informática interactiva en docenas de lenguajes de programación. Matlab es un software con licencia de pago al que no todo el mundo tiene acceso. Además, mediante Jupyter es posible tener acceso al código simplemente accediendo a una dirección web, sin necesidad de tener que tener instalado ningún programa de programación. Esto no es posible con Matlab.
- Curva de aprendizaje. Matlab es un programa utilizado en varias asignaturas impartidas tanto en el Grado de Ingeniería de Telecomunicación como en el Máster. Por otro lado, los códigos implementados en Jupyter están pensados para ser implementados mayormente en Python.
- Complejidad de programación. Dado que la mayoría de los algoritmos están desarrollados en C++, Java o Python, será necesario adaptar muchos de los códigos a Python o a Matlab. Debido a que el uso del lenguaje de programación Python es cada vez más común, el adaptar los códigos a Python será menos laborioso y más útil a futuro.

En la siguiente tabla se exponen los valores de ponderación con los que se calibrará cada criterio de selección:

Criterio	%
Licencia	50
Curva de aprendizaje	25
Complejidad de programación	25

Tabla 2: Peso de los criterios para la selección de la plataforma de trabajo

Y la Tabla 3 presenta la comparativa de cada alternativa:

Criterio	Licencia	Curva de aprendizaje	Complejidad de programación	TOTAL
Jupyter	10	5	6	7,75
Matlab	5	6	5	5,25

Tabla 3: Ponderación de la selección de la plataforma de trabajo

6.3. Proyectos que adaptar

Por otro lado, también se ha hecho un estudio de los diferentes proyectos desarrollados o en fase de desarrollo por Tecnalía para ver cuál de ellos se adaptará a la herramienta que se está desarrollando.

Debido a que el alumno en práctica dispone de un tiempo reducido para realizar el TFM, se han escogido los 4 proyectos que mejor ilustran el trabajo realizado en la división de DEMA. Para ello, será necesario terminar los algoritmos, verificar su correcto funcionamiento y validar tanto los resultados como los datos obtenidos de la base de datos. Además, también será necesario adaptar los resultados para que puedan ser mostrados en una forma clara e interactiva.

Teniendo en cuenta el tiempo estipulado para el trabajo fin de máster, y el trabajo requerido para adaptar cada uno de los algoritmos. De los 6 proyectos planteados inicialmente, se han escogido 4 para adaptar a la herramienta:

- [Zonas multitérmicas de un edificio terciario](#)

Dadas las temperaturas de las diferentes zonas de un edificio terciario, mediante técnicas de analítica de datos, se realiza una clasificación y posterior agrupación de las zonas que tengan un comportamiento y/o un rango de temperaturas similar. Para ellos se hará uso de un algoritmo de procesamiento desarrollado por investigadores de Tecnalia que se centra en trabajar con la magnitud, en este caso la temperatura, que se quiere caracterizar.

Esta práctica dará la opción al alumno de desarrollar su propio algoritmo de procesamiento donde, aplicando y comparando diferentes algoritmos de aprendizaje máquina, podrá comparar los resultados frente a los obtenidos con el algoritmo ya desarrollado y ver la bondad de ambos métodos.

- [Previsión de carga de energía](#)

Se aplicarán métodos regresivos para predecir la demanda de un día en un edificio terciario. Para ello, se obtendrán de la base de datos los datos históricos de dicho edificio. Se podrá también, con el fin de mejorar la predicción, tener en cuenta la temperatura predicha para ese día.

- [Optimizador de energía de próximos 24 horas](#)

Algoritmos de optimización energética que calcularán las horas de encendido y apagado, la temperatura interior estimada, y los horarios de ventilación mecánica de un edificio para las próximas 24 horas con el objetivo de maximizar la eficiencia energética y reducir el consumo energético, manteniendo siempre la temperatura interior de edificio dentro de los márgenes de confort establecidos. A diferencia del caso anterior, en este algoritmo se aplicarán métodos autoregresivos para hacer la estimación de la demanda. Es decir, los resultados obtenidos dependerán de los propios valores anteriores.

Los algoritmos diseñados permitirán fijar ciertos valores de entrada, y ajustándose a ellos, hará una predicción del funcionamiento del sistema que maximice el ahorro energético.

- [Despacho económico de la bomba de calor](#)

Se basa en la aplicación métodos autoregresivos en sistemas HVAC para predecir la demanda y temperatura de un edificio terciario. En base a las horas en las que se quiera estar en confort (parámetro de entrada), propondrá un horario de encendido/apagado de sistema que proporcione un ahorro energético.

Para el procesamiento de esta práctica se hará uso tanto de las temperaturas del interior de edificio como de las temperaturas de impulsión y retorno de la bomba de calor.

- [Sistema de diagnóstico de ineficiencias energéticas en industria](#)

Se basa en el desarrollo de un algoritmo para el sector industrial que, minimizando el coste energético, maximizará la producción y será capaz de detectar las ineficiencias energéticas de máquinas de la industria 4.0.

- [Gestión de la energía en plantas fotovoltaicas](#)

Se basa en aplicar métodos de analítica de datos para hacer una estimación del consumo energético de una planta que disponga placas fotovoltaicas. Conociendo las tarifas, la producción y el consumo horario de la planta, se hará un cálculo para hacer un uso eficiente de la energía fotovoltaica almacenada en las baterías.

Esta herramienta proporciona un ahorro energético ya que, haciendo uso de la energía fotovoltaica que anteriormente ha sido almacenada en las baterías, será capaz de evitar las penalizaciones debidas a los excedentes de potencia.

6.4. Algoritmos de analítica de datos

Para el desarrollo de el algoritmo, se han estudiado los diferentes métodos de agrupación y clasificación de datos disponibles en Python. La librería Scikit-Learn [22] y

SciPy [23] de Python son librerías que disponen de mecanismos de Maching Learning para Python y que proporcionan métodos de analítica de datos y data mining.

La librería Scikit-Learning proporciona métodos de:

- Clasificación: Con diferentes técnicas para identificar a que categoría pertenece un objeto.
- Regresión: Para predecir un atributo de valores continuos asociado a un objeto.
- Clusterización: Agrupación de objetos similares en diferentes grupos.
- Selección de modelos: Técnicas para la comparación, validación y selección de parámetros y modelos.
- Reducción de dimensión: Reducción del número de las variables aleatorias a considerar.
- Preprocesamiento: Extracción de diferentes parámetros a considerar y normalización.

SciPy es un software de código abierto para operaciones matemáticas, científicas y de ingeniería. Entre sus principales módulos destacan:

- Integración
- Optimización
- Operaciones estadísticas
- Clusterización
- Procesado de imágenes
- Procesado de señales

En vistas al algoritmo que se va a desarrollar, se han analizado las diferentes técnicas de clasificación y clusterización disponibles en cada una de las librerías. Una vez estudiadas las alternativas que ofrece cada una, y en base a los objetivos fijados, se han escogido las técnicas que puedan ser útiles para el desarrollo del algoritmo a implementar.

Aunque existen un gran número de técnicas, debido a que se tomará la temperatura como parámetro de entrada, se han seleccionado aquellas que se basan en la distancia entre los diferentes puntos para asignar los grupos. Se han aplicado las técnicas en los

datos a simular y se han estudiado los resultados obtenidos para verificar que se aplican los criterios adecuados.

6.5. Alternativas escogidas

En la Tabla 4 se muestra la elección de las diferentes opciones escogidas para que el trabajo cumpla con los requisitos y objetivos establecidos. La selección se ha realizado en base a los resultados obtenidos al ponderar cada una de las alternativas descritas previamente.

Alternativas Escogidas	
Obtención de datos	Única Base de Datos, EfiDB
Software	Jupyter
Proyectos	<p>Proyecto 1: Zonas multitérmicas de un edificio terciario</p> <p>Proyecto 2: Previsión de carga de energía.</p> <p>Proyecto 3: Despacho económico de la bomba de calor.</p> <p>Proyecto 4: Sistema de diagnóstico de ineficiencias energéticas en industria</p> <p>Técnicas de clusterización Librerías SciPy y Sklearn de Python:</p>
Métodos de Algoritmia	<ul style="list-style-type: none"> • Fcluster • K-Means • Agglomerative clustering

Tabla 4: Alternativas Escogidas para el desarrollo de la herramienta

7. Descripción de las Tareas

En el siguiente apartado se describe brevemente cual ha sido la planificación de tareas que se ha seguido y los recursos hardware y software utilizados para la realización proyecto.

7.1. Recursos Utilizados

7.1.1. Recursos Humanos

En cuanto a los recursos humanos el grupo de trabajo ha sido compuesto por 3 personas. En la Tabla 5 se indica el puesto y responsabilidades de cada uno.

Id	Nombre	Responsabilidad
IS1	Eugenio Perea	Ingeniero Superior & Director de TFM
IS2	Borja Tellado	Ingeniero Superior
IJ	Laura Lázaro	Ingeniero Junio

Tabla 5: Recursos Humanos

7.1.2. Recursos Materiales

En los recursos materiales utilizados se han tenido en cuenta los recursos utilizados tanto para la formación del Ingeniero Junior (IJ) como las herramientas informáticas utilizadas para el procesado de los datos y la documentación y el material de oficina.

No se tendrán en cuenta los recursos utilizados para la obtención y grabación de los datos. Estos aparatos (mayormente sensores) serán utilizados para la recopilación de datos de los diferentes proyectos, los cuales serán volcados a la base de datos. La compra y gestión de los mismos serán responsabilidad de las empresas contratantes.

Recursos Materiales
Ordenador Portátil (x2)
Licencia Microsoft Office
Licencia Windows 10
Material de Oficina

Tabla 6: Recursos Materiales

7.2. Descripción de las Fases y Tareas. Gantt

Fase 0: Gestión y Seguimiento del TFM

Con el fin de asegurar que el Trabajo Fin de Máster se ajusta a las especificaciones de los directores y que cumple con los requisitos y objetivos fijados por los gestores de Tecnalia, se llevará un seguimiento continuo del TFM, desde el inicio hasta que este finalice. En esta tarea se incluirá la planificación y el control del TFM.

Fase 1: Curso de Formación

Fase 1: Curso de Formación		RRHH	Duración (h)	Duración (Sem.)
T1.1	Curso de formación de conceptos básicos del área de trabajo	IJ, IS1	35 (IJ), 30 (IS1)	5
T1.2	Curso de formación de conceptos básicos de Python	IJ, IS2	75 (IJ), 50 (IS2)	6
T1.3	Cursos de formación de Analítica de datos	IJ, IS2	60 (IJ), 30 (IS2)	6

Tabla 7: Fase1: Curso de formación

En esta fase el alumno obtendrá la formación necesaria para llevar a cabo el Trabajo Fin de Máster y empezará a familiarizarse con los términos y proyectos que se llevan a cabo en la División de Energía y Medio Ambiente.

T1.1: Curso de formación de conceptos básicos del área de trabajo. Se impartirán unos breves cursillo al alumno con el fin de que se familiarice con algunos de los conceptos básicos de área de la eficiencia energética. También se le facilitará documentación e información de lectura.

T1.2: Curso de formación de conceptos básicos de Python. Debido a que el alumno no tiene experiencia previa con el lenguaje de programación Python, será necesaria una breve formación antes de empezar a desarrollar la herramienta.

T1.3 Cursos de formación de Analítica de datos. Inicialización del alumno en el manejo de la analítica de datos.

Fase 2: Documentación

Fase 2: Documentación		RRHH	Duración (h)	Duración (Sem.)
T2.1	Selección de los demostradores a realizar	IS1	30	4
T2.2	Documentación de cada una de las prácticas	IJ	40	5

Tabla 8: Fase 2: Documentación

T2.1: Selección de los demostradores a realizar. Tras un estudio de los diferentes proyectos llevados a cabo en la división, se decidirá cuales serán aquellos que mejor engloben las capacidades con las que se trabaja en DEMA. Esta será una responsabilidad que se llevará a cabo por el director del proyecto fin de Máster.

T2.2: Documentación de cada una de las prácticas. Para poder desarrollar los escenarios de la mejor manera posible, será necesario que el ingeniero junior se familiarice con cada uno de los proyectos escogidos.

Fase 3: Estudio, Selección y Adaptación

Fase 3: Estudio, Selección y Adaptación		RRHH	Duración (h)	Duración (Sem.)
T3.1	Estudio de las diferentes bases de datos disponibles	IJ, IS2	14, 14	4
H3.1	Diseño de la Base de Datos			
T3.2	Adaptación de los parámetros	IJ	25	4
T3.3	Estudio de los diferentes algoritmos	IJ, IS2	15, 15	6
H3.2	Selección de los algoritmos a implementar			

Tabla 9: Fase 3: Estudio, selección y adaptación

T3.1: Estudio de las diferentes bases de datos disponibles. Para empezar con el desarrollo de las prácticas, será necesario, en primer lugar, analizar las diferentes bases de datos con las que se pueda trabajar. De esta forma se intentará construir una única base de datos capaz de recoger, de forma ordenada, los datos guardados en el resto de bases de datos (H3.1).

T3.2: Adaptación de los parámetros. Se adaptarán y adecuarán los diferentes parámetros de las bases de datos para que sea posible recogerlos todos en una única base de datos centralizada. También se tendrá que tener en cuenta que estos datos deberán ser manejables haciendo uso del lenguaje de programación Python.

T3.3: Estudio de los diferentes algoritmos. De igual manera se estudiarán y elegirán (H3.2) las diferentes técnicas de analítica de datos y algoritmos con los que trabajará el Ingeniero Junior para desarrollar el algoritmo.

Fase 4: Diseño y Desarrollo de la Herramienta

Fase 4: Diseño y Desarrollo de la Herramienta		RRHH	Duración (h)	Duración (Sem.)
T4.1	Diseño y desarrollo de la herramienta en el entorno Jupyter	IJ, IS1, IS2	300, 30, 30	16
T4.2	Desarrollo del demostrados Zona Multitérmica	IJ, IS1, IS2	125, 25, 25	10

Tabla 10: Fase 4: Diseño y desarrollo de la herramienta

T4.1: Diseño y desarrollo de la herramienta en el entorno Jupyter. Una vez escogida la base de datos y los algoritmos que se van a utilizar se procederá a diseñar la interfaz de la que harán uso los trabajadores de Tecnalía para exponer los diferentes proyectos. Para ello, será necesario adaptar los diferentes algoritmos y escenarios.

Se programará cada una de las prácticas en función de los algoritmos que se aplicarán. Cada uno de los escenarios tendrá interfaz gráfica diferente, diseñadas con flexibilidad para poder dar al usuario la opción de escoger entre varios parámetros de entrada y así poder ver cómo afecta a los resultados.

T4.2: Desarrollo del demostrados Zona Multitérmca. Se desarrollará e implementará un nuevo algoritmo que, haciendo uso de técnicas de machine learning y analítica de datos, realizará agrupaciones en función de comportamiento de diferentes zonas de un edificio. Se incluirá un nuevo Notebook de Jupyter con el nuevo algoritmo.

Fase 5: Simulaciones

Fase 5: Simulaciones		RRHH	Duración (h)	Duración (Sem.)
T5.1	Simulaciones y verificación de correcto funcionamiento	IJ, IS1	30, 15	5
T5.2	Cambios y solución de errores	IJ, IS1	25, 10	2
H5	Verificación de correcto funcionamiento			

Tabla 11: Fase 5: Simulaciones

T5.1: Simulaciones y verificación de correcto funcionamiento. Para comprobar que las herramientas diseñadas funcionan correctamente, se harán varias simulaciones y análisis de cada interfaz variando los parámetros de entrada.

T5.2: Cambios y solución de errores. En el caso de que encuentre algún fallo o de que el resultado no sea el esperado, se revisará y solucionará el problema.

Fase 6: Redacción del Informe

Fase 6: Redacción del Informe		RRHH	Duración (h)	Duración (Sem.)
T6.1	Redacción del informe	IJ	40	4
T6.2	Aprobación de los directores del TFM	IS1, IS2	8, 8	1
H6	Entrega del TFM			

Tabla 12: Fase 6: Redacción del informe

T6.1: Redacción del informe. Para dar por concluido el trabajo, será necesario redactar un informe donde se explique de forma detallada el trabajo llevado a cabo y el funcionamiento de cada una de las interfaces desarrolladas. Aunque algunos puntos del informe se irán redactando de forma paralela al desarrollo del TFM, el documento se finalizará al finalizar las interfaces y el algoritmo, una vez se hayan obtenido los resultados.

T6.2: Aprobación de los directores del TFM. Este informe será revisado y aprobado por los directores del trabajo antes de dar por finalizado el TFM (**H6**).

Duración total del trabajo

El desarrollo del presente TFM fue iniciado el 7 de septiembre del 2017 y se finalizado para el 13 de agosto del 2018. Hay que destacar que el alumno, al estar contratado por la empresa Tecnalía, ha dedicado tiempo a actividades adicionales de la empresa que no son parte del desarrollo y redacción del presente TFM.

Como se muestra en la siguiente ilustración, se ha estimado que la duración total del proyecto será de aproximadamente 11 meses:

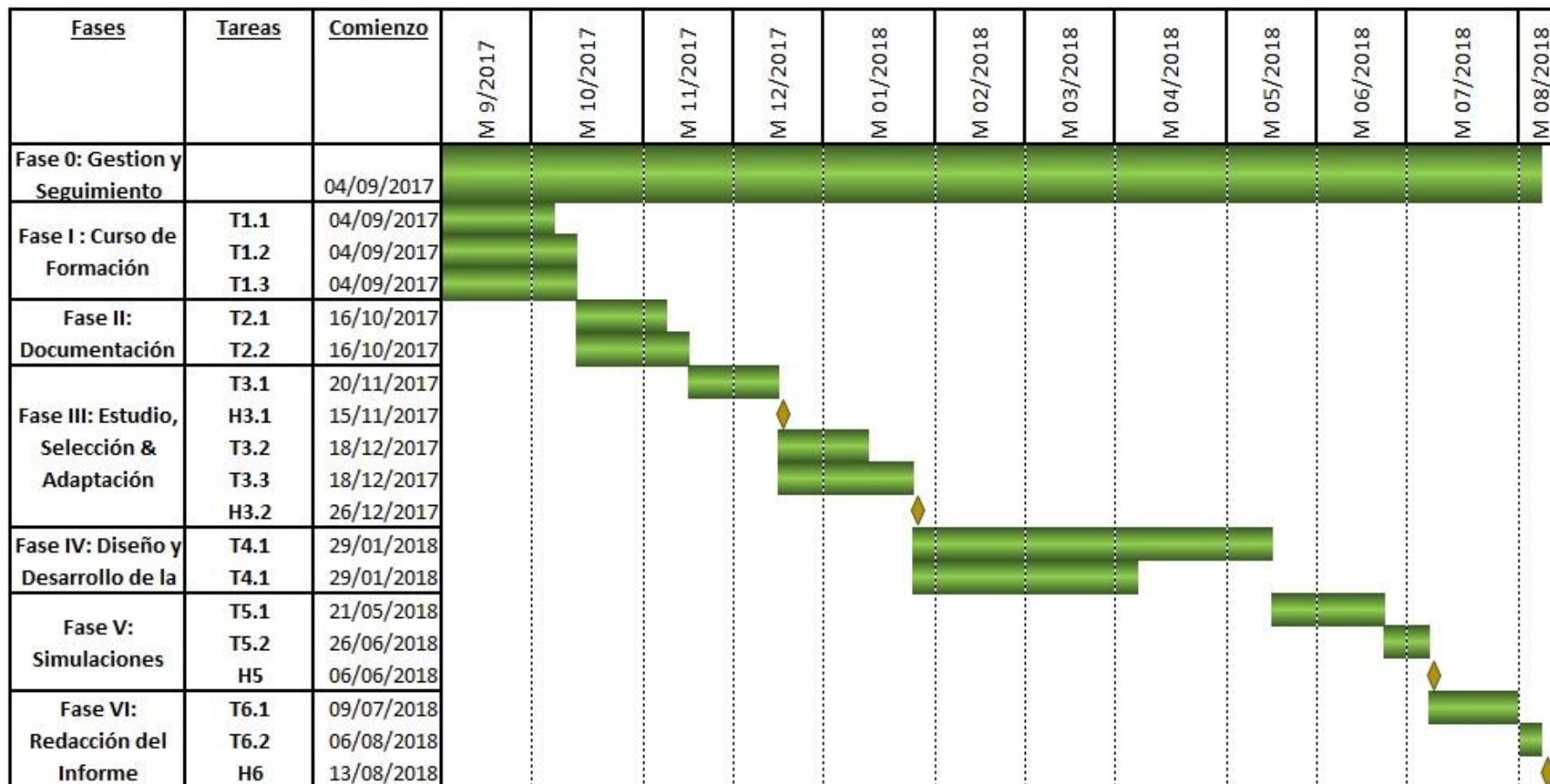


Ilustración 19: Gant. Tiempo de dedicación previsto para las diferentes tareas y fases del TFM

8. Metodología

En este apartado, se va a describir la metodología que se ha seguido para realizar los diferentes escenarios y desarrollar el nuevo algoritmo; para ello se explicarán los pasos que se han seguido y las simulaciones que se han ido realizando con el fin de verificar el correcto funcionamiento de los mismos.

8.1. Diseño de la herramienta

En la Ilustración 20 se muestran los pasos seguidos para el diseño de los diferentes escenarios de la herramienta que se va a diseñar:

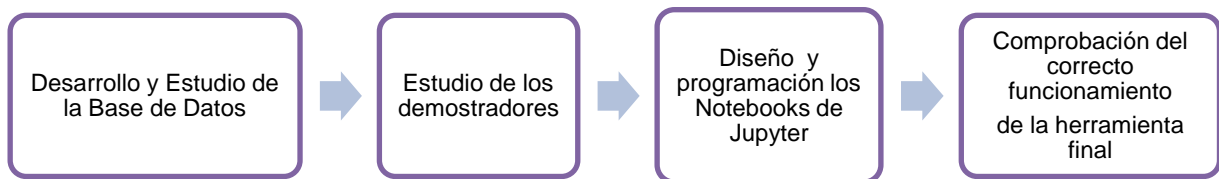


Ilustración 20: Pasos seguidos para la realización de cada práctica

8.1.1. Base de datos

Para el desarrollo de la herramienta, en primer lugar, ha sido necesario recopilar los datos de los diferentes edificios terciarios y hogares que se utilizarán para testear los diferentes algoritmos desarrollados. Los sensores de captación de datos que se utilizarán en cada uno de los edificios estarán escogidos por cada uno de los propietarios y Tecnia solo tendrá acceso a los datos medidos.

Aunque habrá escenarios que no hagan uso de la base de datos, para disponer de todos los datos en una única fuente, se ha diseñado una base de datos centralizada que será capaz de guardar todos los datos de una forma bien estructurada (ANEXO I). Cada elemento de medición dispondrá de un identificador que lo identificará unívocamente. El identificador permitirá relacionarlo con el resto de elementos del edificio y con el algoritmo a implementar. Una vez se haya diseñado la estructura de la base de datos, se volcarán los datos en la misma.

Antes de empezar a trabajar con los algoritmos, se realizarán algunas pruebas con el fin de comprobar que los datos se estén tomando correctamente e identificar anomalías en la captación o en el valor de los datos.

8.1.2. Demostradores

Por otro lado, se ha procedido a adecuar los algoritmos seleccionados para los diferentes demostradores de la herramienta. De esta forma, tomando los datos de la base de datos, y una vez establecidos los diferentes parámetros de entrada, se adecuarán los algoritmos para obtener unos resultados que ilustren el modo de funcionamiento de cada uno de ellos. Como serán simulaciones que se realizarán a modo de apoyo en las exposiciones, será necesario asegurarse del correcto funcionamiento del algoritmo y de que el tiempo de simulación no es demasiado elevado.

Aunque se han utilizado algoritmos que ya estaban desarrollados, se aprovechará el desarrollo de la herramienta para revisarlos y añadir mejoras. Además, también será necesario adecuar los diferentes códigos de forma de que todos sigan una estructura similar. Algunos de los códigos han sido desarrollados en diferentes lenguajes como C++ o Java, por lo que será necesario traducirlos a Python en primer lugar.

En vistas a diseñar una herramienta interactiva que permita al usuario variar los parámetros de entrada, será necesario identificar para cada caso cuáles serán los parámetros de entrada que puedan tener repercusión en los resultados de salida.

Para cada caso, se testeará cada algoritmo con los datos de la base de datos para cerciorar que los datos que se están utilizando son los correctos y de que los resultados obtenidos son los esperados (Ilustración 21).

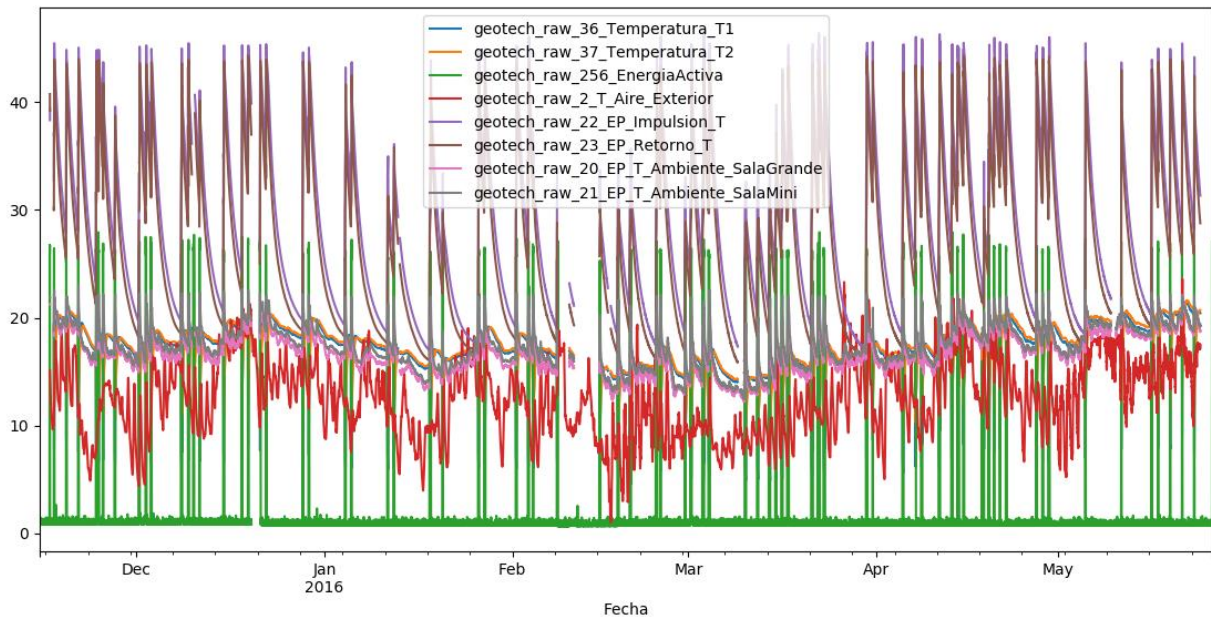


Ilustración 21: Datos del Museo Btek, cargados de la base de datos

8.1.3. Interfaz Gráfica

Una vez que definida la base de datos y teniendo seleccionado los algoritmos que vamos a implementar, se desarrollarán los Notebook de Jupyter para cada uno de los escenarios. El Jupyter Notebook es una aplicación web de código abierto que permite crear y compartir documentos que contengan códigos, ecuaciones, visualizaciones y textos en vivo.

Aunque en cada caso serán necesarios diferentes parámetros de entrada, todos los Notebooks diseñarán siguiendo una misma estructura con el propósito de simplificar el aprendizaje del usuario que vaya a simular la herramienta y de ayudar al exponente al seguir la explicación. Aunque a simple vista todos los escenarios tendrán una apariencia parecida, cada uno llevará el código correspondiente un proyecto específico.

En la siguiente imagen se muestra una interfaz de uno de los Notebooks desarrollados. Como ya se ha mencionado, todos los Notebooks serán diseñados en base a un mismo esqueleto. Se ha añadido anexo al documento un Manual de Usuario (ANEXO II) donde se explica de forma detallada la forma en la que manejar la herramienta.

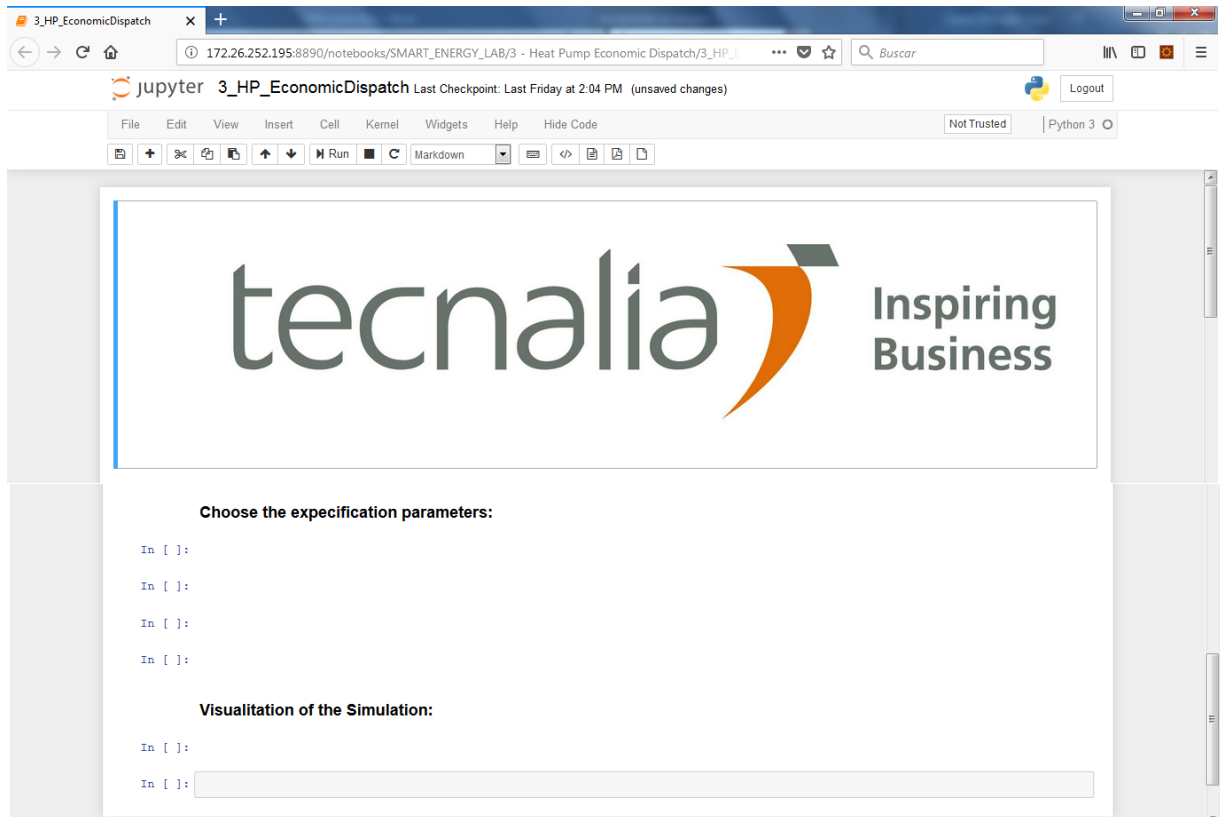


Ilustración 22: Estructura Interfaz Gráfica Jupyter

Con el fin de poder hacer un estudio de los resultados obtenidos, el usuario tendrá la opción de guardar los resultados a forma de imagen o fichero .csv, en función del escenario simulado.

8.1.4. Máquina Virtual

Aunque hoy en día Python es uno de los lenguajes de programación más utilizado para el desarrollo de algoritmos, no todo el mundo dispone del software requerido para ejecutar programas desarrollados en este lenguaje en su PC. Además, el hecho de que cada poco tiempo haya disponibles versiones nuevas del software puede suponer que el código desarrollado no sea compatible con la versión de Python disponible en el PC de la persona que hará uso de la herramienta desarrollada.

Al ser el Notebooks de Jupyter una aplicación web, se ha hecho uso de una máquina virtual de TecNALIA situada en la DMZ para asegurarnos de que todos los usuarios que accedan a través de la red de TecNALIA puedan tener acceso a la versión actualizada de la herramienta.

En la Ilustración 23 se muestra un esquema de una DMZ. La DMZ o zona desmilitarizada [24] es una zona insegura de acceso a determinados equipos que se ubica entre la red interna de la organización de Tecnalía e Internet. Las conexiones desde la red interna y la externa (Internet) están permitidas, pero en general, las conexiones de la DMZ solo se permiten a la red externa. Es una forma de proteger la red interna de diferentes intrusos que puedan poner en riesgo la seguridad de los equipos. Es una forma de evitar posibles ataques a la red interna de la empresa.

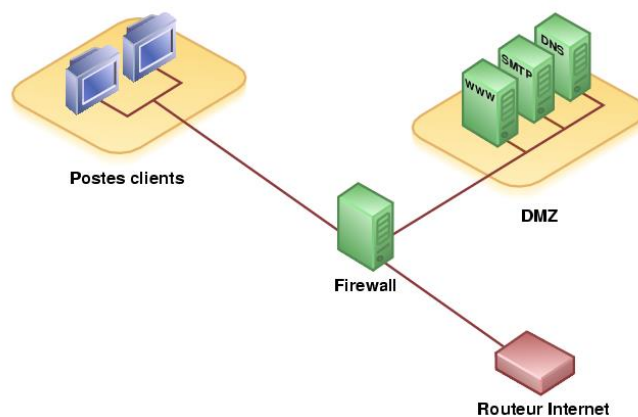


Ilustración 23: Esquema Simplificado de una DMZ

8.2. Diseño del algoritmo

En el siguiente diagrama de bloques se muestra los pasos seguidos para el diseño del algoritmo:

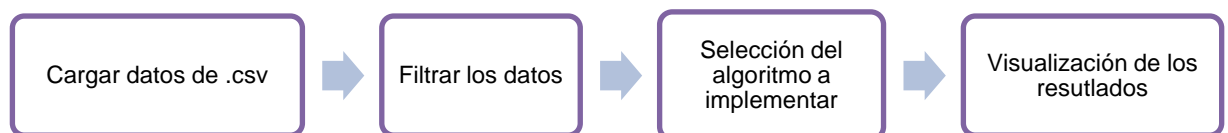


Ilustración 24: Pasos seguidos para el diseño del algoritmo

Para el diseño del algoritmo, en primer lugar, será necesario volcar los datos recopilados en la herramienta que se va a utilizar. En este caso, se trabajará con en el entorno Python. No será necesaria la validación de los datos ya que se utilizarán los mismos datos que los utilizados en uno de los escenarios desarrollados para la herramienta.

Una vez cargados los datos, se discretizarán en función de los trimestres de año. El hecho de separar por trimestres hará que el rango de temperaturas sea más similar y que los resultados sean, por lo tanto, más precisos. Los datos se separarán de la siguiente manera:

- Invierno: diciembre, enero, febrero.
- Primavera: marzo, abril, mayo.
- Verano: junio, julio, agosto.
- Otoño: septiembre, octubre, noviembre.

Una vez cargados los datos, mediante la técnica de agrupación de K-Means se agruparán los días en función de las horas que está activo el sistema de climatización. De esta manera, se diferenciarán los tipos de días para dar al usuario la opción de escoger entre:

- Jornada Normal: Sistema de climatización encendido 8 horas, por lo general de lunes a jueves.
- Jornada Intensiva: Sistema de climatización encendido 6 horas, por lo general los viernes.
- Fines de semana: Sistema de climatización apagado.

A continuación, se realizará el procesamiento de los datos seleccionados. Se han escogido los métodos en base a los objetivos establecidos para el algoritmo. Entre las diferentes técnicas disponibles, se el usuario tendrá la opción de escoger entre los siguientes:

- Fcluster con criterio de maxclus
- Fcluster con criterio de distancia
- k-Means
- AgglomerativeClustering

Debido a que la el comportamiento de la temperatura en las diferentes zonas térmicas varía en función de que el sistema este encendido o apagado, se separarán los días en tres periodos diferentes: sistema apagado antes del encendido, sistema encendido y sistema apagado después del encendido. El algoritmo realizará una clusterización (con el método escogido) para cada periodo

de cada día y le asignará clúster a cada una de las zonas térmicas. Los resultados de las agrupaciones se irán representando gráficamente y se irán guardando en memoria. Una vez procesados todos los datos, se cargarán los datos guardados y se representará la matriz correlación. La matriz correlación servirá para conocer la relación que existe entre el comportamiento de las diferentes zonas térmicas, para luego agrupar aquellas que se comporten de forma similar.

9. Descripción de la Solución

A continuación, se van a trabajar y describir cada uno de los cuatro demostradores y el nuevo algoritmo que conforman la solución Smart Energy Lab, explicando los puntos fundamentales de los algoritmos implementados.

9.1. Demostrador #1: Zonas multitérmicas de un edificio terciario

El Modelo de Análisis de Zonas Térmicas (TZAM) tendrá como principal objetivo la identificación de las temperaturas características en un edificio terciario dotado de múltiples sensores de temperatura.

Algo importante para alcanzar el confort térmico en un edificio entero es asumir que todo el edificio no puede tener el mismo comportamiento térmico. Para el siguiente escenario, se ha trabajado con un algoritmo que, en función de las temperaturas del interior del edificio y de un rango de confort dado por el usuario, será capaz de agrupar en clústeres las diferentes zonas térmicas. Cada agrupación de zonas térmicas deberá ser administrado por un único punto de referencia, que será la temperatura característica del grupo o clúster.

Se va a trabajar con datos históricos de temperatura de una de las plantas del edificio 700 de Tecnia, en Derio, Bizkaia. Los datos se cargarán de un archivo de tipo csv de nombre *combinedFile.csv* que contendrá las temperaturas de las diferentes zonas térmicas en valores 10 minutarios medidos desde el 20 de febrero del 2013 hasta el 30 de septiembre del 2014.

El comportamiento térmico de un edificio está completamente relacionado con los comportamientos estacionales, que deberán diferenciarse para generar patrones térmicos similares. Por lo tanto, será necesario establecer ciertos parámetros que permitan discretizar los datos de entrada en base al periodo del año y en el tipo de días. Estos umbrales permitirán que el sistema defina diferentes comportamientos en el sistema de climatización para comprender adecuadamente el comportamiento de las diferentes zonas térmicas.

El número K de temperaturas características de un determinado edificio es la agrupación de todas las temperaturas interiores que satisfacen la siguiente expresión:

$$\frac{1}{K} \sum_{k=1}^K \sum_{i=1}^N \frac{|T_i^j - T_{ch}^k|}{N} \leq \Delta_{ch} = T_{conf}^{max} - T_{conf}^{min}$$

La agrupación de las zonas multitérmicas se realizará en base a un número k, dada una estrategia de inicialización y siguiendo un proceso iterativo. Como entrada, este procedimiento necesitará la matriz de temperaturas y un valor adicional que representará la estrategia de inicialización. Este parámetro será un parámetro de usuario que podrá adoptar tres valores diferentes:

- Cool: El proceso comenzará con el análisis de dos zonas más frías. Una vez que se realice este análisis, el proceso continuará en orden ascendente de temperatura.
- Hot: el proceso comenzará con el análisis de dos zonas más calientes. Una vez que se realice este análisis, el proceso continuará en orden descendente de temperatura.
- Random: el proceso elegirá aleatoriamente las zonas que a analizar.

La agrupación de las zonas térmicas en los diferentes grupos se realizará en función de dos parámetros de entrada establecidos por el usuario:

- Máxima diferencia térmica: máxima diferencia térmica que se contempla para que dos zonas térmicas entren dentro de una misma agrupación.
- Número de desviaciones permitidas: el porcentaje de tiempo que se puede desviar la máxima diferencia térmica sin que sea considerada como parte de otro grupo.

El análisis de las zonas se realizará siguiendo este proceso secuencial:

- Usando las temperaturas registradas por dos zonas de estudio a lo largo del tiempo, se creará una línea promedio para todas las marcas de tiempo.
- Para cada marca de tiempo, se verificará la diferencia de temperatura entre el promedio y la zona térmica, si esta diferencia es mayor que T_{conf}^{max} se registrará una desviación, lo que implicará que, en ese momento ambas zonas térmicas no podrán ser manejadas usando el mismo punto de ajuste.
- Una vez que se calcule el número de desviaciones, se podrán darse dos escenarios diferentes:
 - Si el número de desviaciones es menor que las desviaciones permitidas, se supondrá que ambas zonas son parte del mismo clúster y que se pueden gestionar cómodamente utilizando el mismo punto de consigna.

- Si el número de desviaciones es mayor que las desviaciones permitidas, las zonas térmicas serán parte de diferentes grupos. En este sentido, todas las zonas analizadas previamente compondrán un clúster, dejando la nueva zona como parte de una nueva agregación.

Este procedimiento se ejecutará hasta que todas las zonas térmicas del edificio tengan un clúster asignado. Además de obtener como resultado las agrupaciones realizadas, los resultados se representarán en los siguientes gráficos:

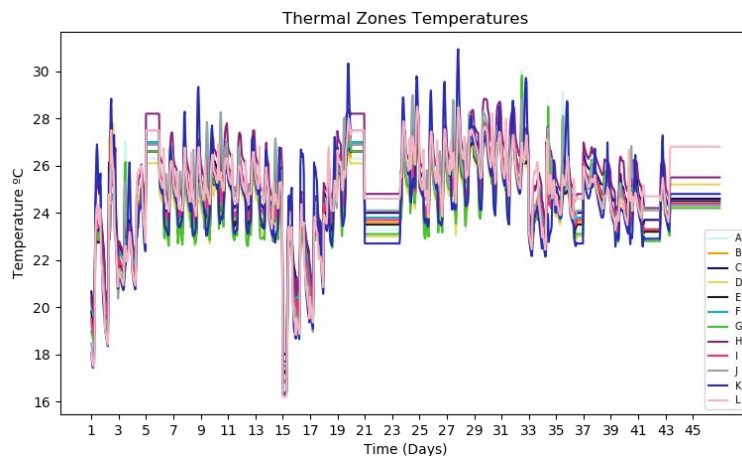


Ilustración 25: Temperaturas para las 12 zonas térmicas

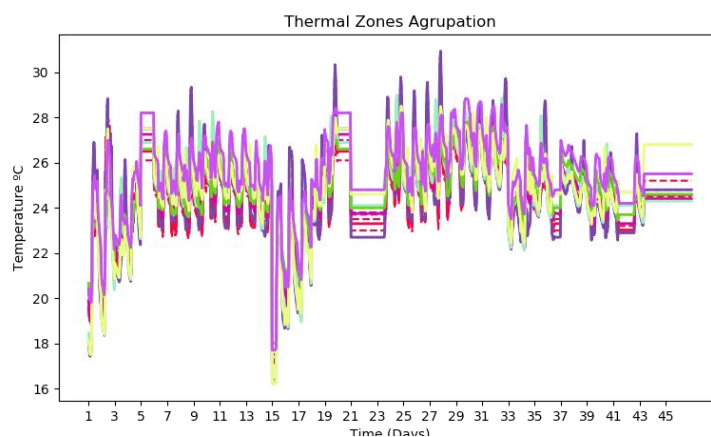


Ilustración 26: Agrupaciones realizadas (con colores) de las diferentes zonas térmicas

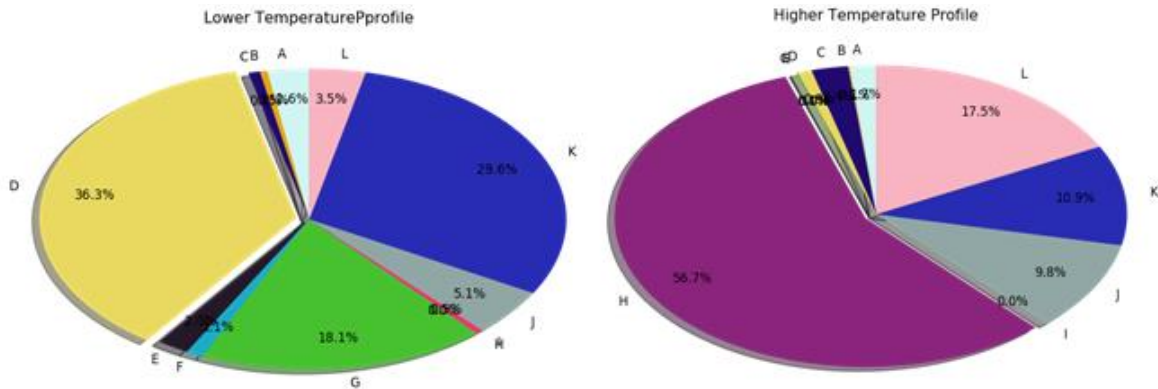


Ilustración 27: Perfiles de Máximos y Mínimos de temperatura (%)

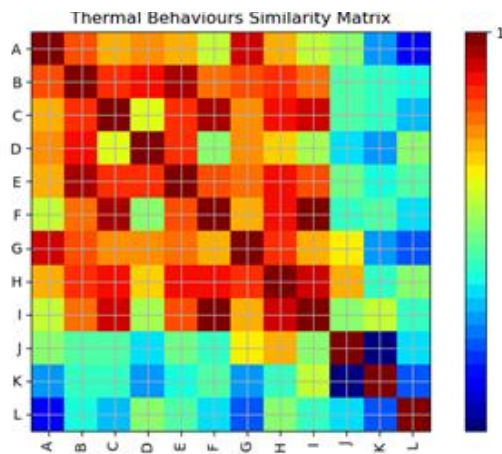


Ilustración 28: Matriz de correlación. Similitud entre el comportamiento entre zonas térmicas

- Ilustración 25: La primera gráfica mostrará las diferentes temperaturas de las diferentes zonas térmicas para todo el espectro temporal, cada una representada con un color diferente.
- Ilustración 26: A continuación, se mostrará las diferentes temperaturas de las diferentes zonas térmicas para todo el espectro temporal. Cada clúster tendrá asignado un color y las zonas térmicas que pertenezcan a ese clúster estarán representadas con ese color.
- Ilustración 27: Se mostrarán también los perfiles de máximos y mínimos de temperatura. Esta gráfica indicará porcentualmente el tiempo que ha sido cada zona la más fría y la más caliente.
- Ilustración 28: Por último, se representará la matriz correlación. La matriz correlación indicará la similitud en el comportamiento entre las diferentes zonas térmicas.

9.2. Demostrador #2. Previsión de carga de energía

El objetivo principal del proyecto de previsión de la demanda de energía es el de realizar una estimación de la generación de energía para el siguiente periodo de tiempo de un edificio en base a los datos históricos de consumo y el consumo real del último período. Con el fin de mejorar la calidad de los resultados y realizar una mejor predicción, el algoritmo desarrollado también dará la opción de tener en cuenta las previsiones meteorológicas del día a simular.

Las entradas necesarias para pronosticar la generación de energía serán las siguientes:

- **Demanda energética:** La información proporcionada tendrá el objetivo de ayudar al algoritmo a inferir la relación entre el clima en un área determinada y el consumo de energía en dicha área.
- **Datos históricos:** Datos de temperatura medidos en el edificio en fechas anteriores.
- **Variables climáticas:** En el caso de que el pronóstico del tiempo sea considerado como entrada, también deberán considerarse los datos históricos del pronóstico del tiempo.

La herramienta realizará una correlación entre los datos históricos y la demanda mediante técnicas autoregresivas. De esta manera, aprenderá de los casos pasados para que realice una buena estimación de los casos futuros. Por lo tanto, cuanto mayor sea la cantidad de datos históricos que se posean, mejor será la capacidad de aprendizaje del algoritmo, debido a que tendrá un mayor número de registros con los que poder comparar y extrapolar los resultados a casos futuros.

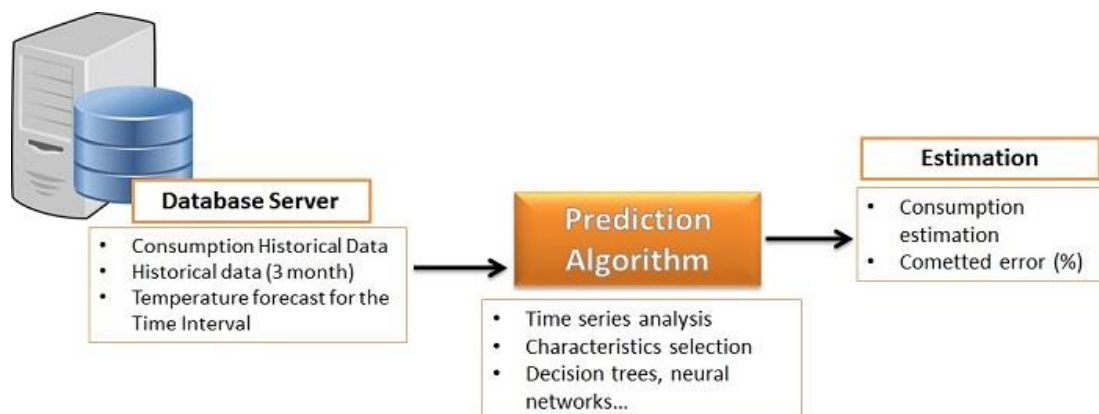


Ilustración 29: Diagrama de bloques del Escenario 2

El algoritmo estará compuesto por los siguientes módulos:

- Automatización de los datos de entrada: Dependiendo del modelo de negocio con el que se trabaje, el cliente podrá proveer los datos en varios formatos provenientes de diferentes modelos SCADA. La automatización de los datos necesitará cierto trabajo de ingeniería.
- Limpieza de datos: Para evitar que haya datos nulos o valores invalidados en el set de datos que se utilice como entrada al algoritmo, se realizará una limpieza de los datos de entrada.
- Selección de características: No todas las magnitudes de entrada estarán correlacionadas con la demanda energética. Será importante seleccionar y trabajar con aquellas que tengan relación con la demanda energética.
- Entrenamiento: Para que la estimación realizada sea válida, será necesario al menos un mes de datos históricos por cada estación. El modelo auto-regresivo se entrenará a partir de los datos históricos para establecer el coeficiente interno correcto para el problema a pronosticar. El entrenamiento se realizará periódicamente mediante una ventana deslizante. Una vez recopilados los datos, el algoritmo entrenará el modelo antes de su ejecución para obtener los datos históricos más recientes.
- Visualización: Se visualizarán con dos líneas sobrepuestas los resultados predichos frente a los reales para el día seleccionado. Además, también se indicará el valor del error medio horario y del error medio diario cometido en la estimación realizada (Ilustración 30).

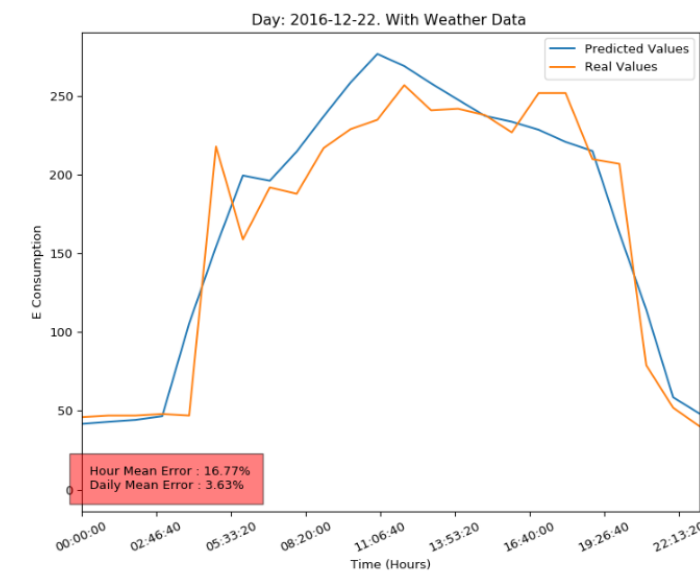


Ilustración 30: Resultados para la simulación del escenario 2

9.3.Demostrador #3. Despacho económico de la bomba de calor

El consumo de energía depende de múltiples factores internos y externos, condiciones climáticas, temperatura del suministro de entrada, Coeficiente de rendimiento (COP), aislamiento térmico del edificio, ocupación, etc. Para este escenario, se trabajará con un modelo de caja negra basado en métodos de inteligencia artificial para simular el modelado del edificio. Para ello, se recopilarán datos del museo BTEK. BTEK es un edificio situado en el parque tecnológico de Derio, Bizkaia, el cual dispone de una bomba de calor y varios sensores de medida. Para el modelado del edificio se trabajará tanto con las temperaturas medidas del edificio como con las temperaturas de impulsión y retorno de la bomba de calor.

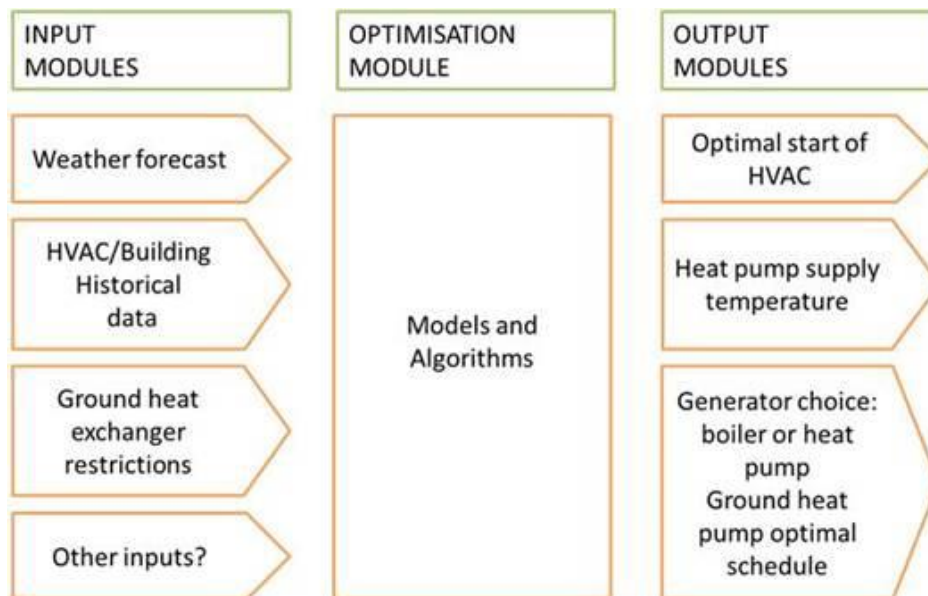


Ilustración 31: Diagrama de bloques del escenario 3

Mediante el desarrollo de este algoritmo se busca minimizar el consumo de energía del edificio, ajustando las secuencias de tiempo ON / OFF del sistema de climatización y garantizando las condiciones de confort impuestas.

El esquema de caja negra propuesto se implementará a partir de esquemas de regresión artificiales que extraerán la relación no lineal entre las variables de entrada y salida para aprender de los datos experimentales. En este caso, las magnitudes BTEK monitoreadas constituirán el conjunto de datos de entrenamiento. Las condiciones de confort del edificio y la hora de encendido y apagado del sistema vendrán, o bien establecidas por el usuario, o se tomarán de los valores históricos reales que se establecieron en el día concreto que se desee simular.

Como se muestra en la Ilustración 32, serán necesarios los siguientes módulos como parámetros de entrada al algoritmo:

- Módulo de pronóstico del tiempo: magnitudes del clima para las próximas 24 horas.
- Calendario con valores de temperatura interior y las restricciones de confort impuestas por el usuario.
- Magnitudes históricas medidas para el entrenamiento de los modelos de caja negra implementados como edificio térmico.

Las primeras dos medidas (comienzo del día) se usarán como semilla inicial para comenzar el esquema del modelo de control predictivo (MCP) recursivo propuesto. De esta forma, el esquema de MCP propuesto determinará los tiempos On/Off que asegurarán las condiciones de confort al tiempo que reducirán las horas de operación. Además, dependiendo de la previsión de temperatura y del rango de confort exigido por el usuario, el sistema reconocerá de forma autónoma el modo de funcionamiento de la bomba (modo frío o modo calor).

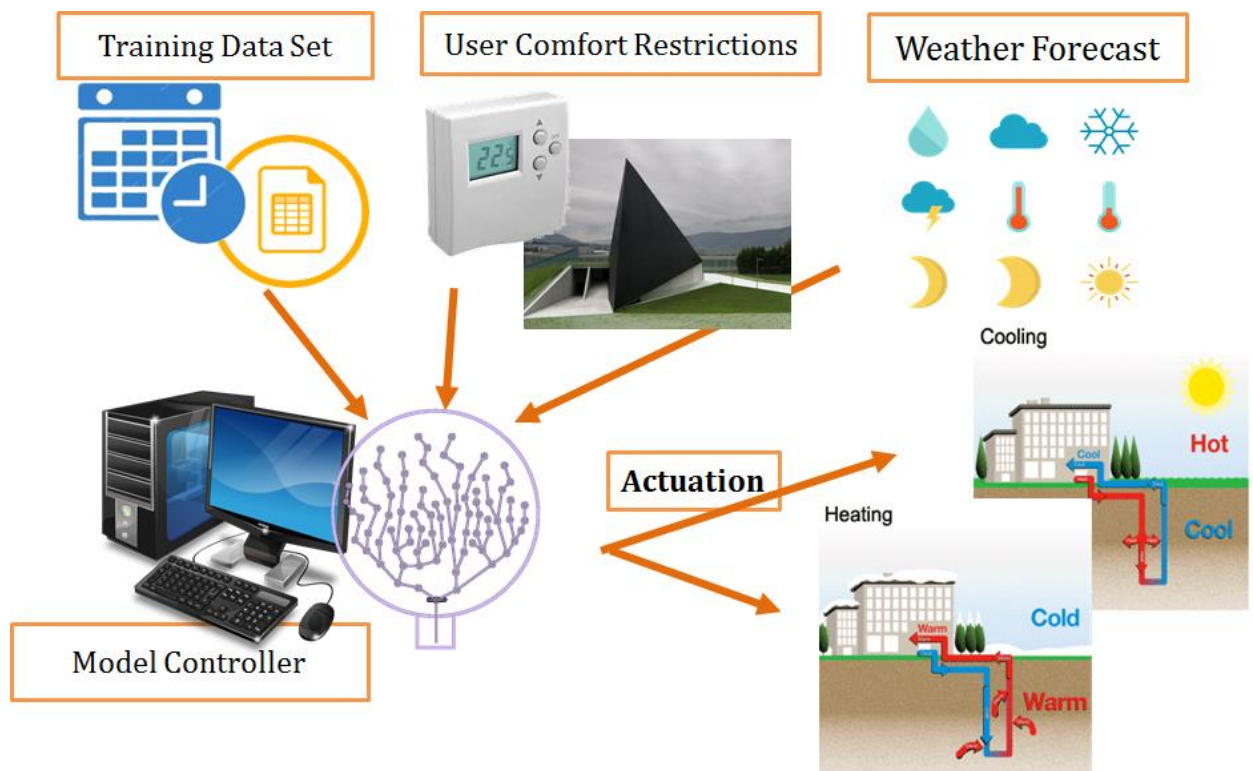


Ilustración 32: Esquema del de Modelo de Control Predictivo

En la siguiente imagen se muestran los resultados obtenidos al ejecutar la simulación para el día 25 de noviembre del 2015. Los parámetros de simulación introducían exigían estar en confort entre las 11:15 y las 14:45 horas del día simulado. Como se muestra en la imagen, la simulación calcula que el sistema debería de encenderse a las 10:00 y apagarse a las 13:15 para cumplir con las exigencias requeridas por el usuario. De esta forma se consigue acortar el tiempo que el sistema estará encendido, lo que se traduce en un ahorro energético y económico.

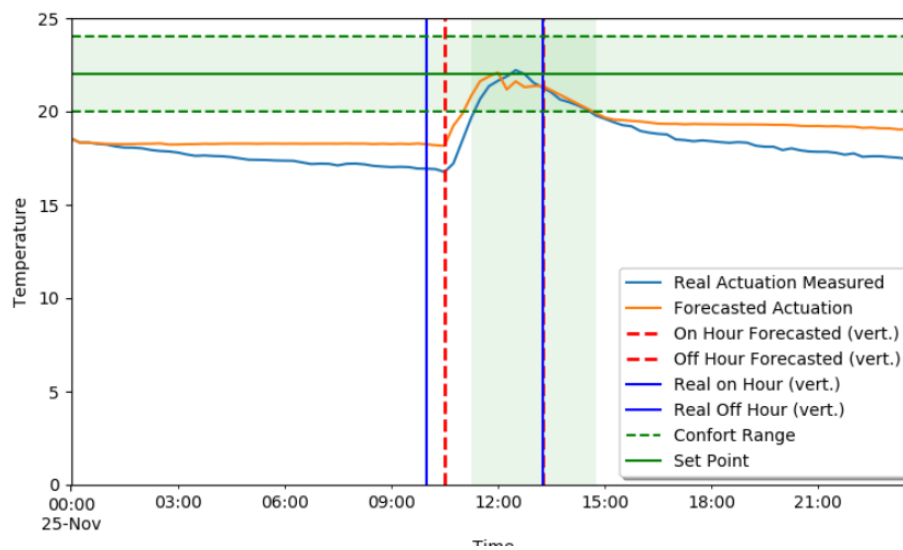


Ilustración 33: Resultados de simulación escenario 3

9.4. Demostrador #4. Sistema de diagnóstico de ineficiencias energéticas en industria

La mejora en la gestión de la energía supone para las empresas un ahorro energético y de costo a través de una toma de decisiones informada. Con la implementación de prácticas de ahorro de energía para plantas, líneas, procesos, máquinas y operaciones se ha desarrollado un algoritmo que, mediante la evaluación de los datos recabados de diferentes máquinas de la planta será capaz de identificar los días con ineficiencias energéticas y en base al tipo de día, lo comparará con otros días similares donde el ahorro energético haya sido mayor.

Dado un conjunto de medidas, el algoritmo extraerá los perfiles de consumo característicos de una determinada planta y clasificará cada día en términos de su patrón de consumo tipo. Será necesario

disponer, con una profundidad temporal mínima de menos un mes, de los datos históricos de las máquinas de la empresa que se deseen analizar.

El proceso estará dividido en diferentes funciones:

- Función carga de datos: extraerá los perfiles de consumo.
- Función entrenamiento para generación de patrones: Dado un conjunto de datos de consumos y de producción (de 1 a 3 meses) la herramienta entrenará y aprenderá comportamientos, siendo capaz de generar diferentes patrones o clústeres.
- Función clasificación de nueva curva de carga en el conjunto de patrones: clasificará el consumo de una nueva curva de carga de el intervalo temporal en estudio (configurable por el usuario), en uno de los posibles patrones tipo.
- Función análisis de energía de carga dentro de su patrón: Analizará si dicho consumo respecto de su patrón tipo es estadísticamente razonable o si por el contrario puede suponer un exceso de consumo energético. En esta función se tratará de buscar comportamientos anómalos. Será posible detectar si un consumo excesivamente bajo dentro del patrón está ligado a que algún elemento no haya funcionado correctamente. De ser así, se generará una alarma.
- Función análisis producción vs energía dentro de su patrón: Comparará el consumo energético vs producción de ese día y los de los días pertenecientes al patrón.

Para la función “*entrenamiento para generación de patrones*”, el algoritmo necesitará tener acceso a datos de consumo eléctrico y de producción de la planta con una profundidad temporal de al menos 1 mes. Estos datos servirán como datos de entrenamiento. Haciendo uso de técnicas de agrupación disponibles en la librería Sklearn de Python, se podrán diferenciar los diferentes patrones de comportamiento de los días disponibles en el set de entrenamiento. En la Ilustración 34 se muestra los resultados obtenidos una vez procesados los datos de entrenamiento para una planta industrial de Bizkaia. Como se puede ver, en base al comportamiento temporal, los días serán agrupados en 5 grupos diferentes. A medida que se vayan procesando nuevos datos, el algoritmo marcará de color negro las curvas de carga cuyo comportamiento no se asemeje a ningún grupo y los etiquetará como *outliers* o alarmas.

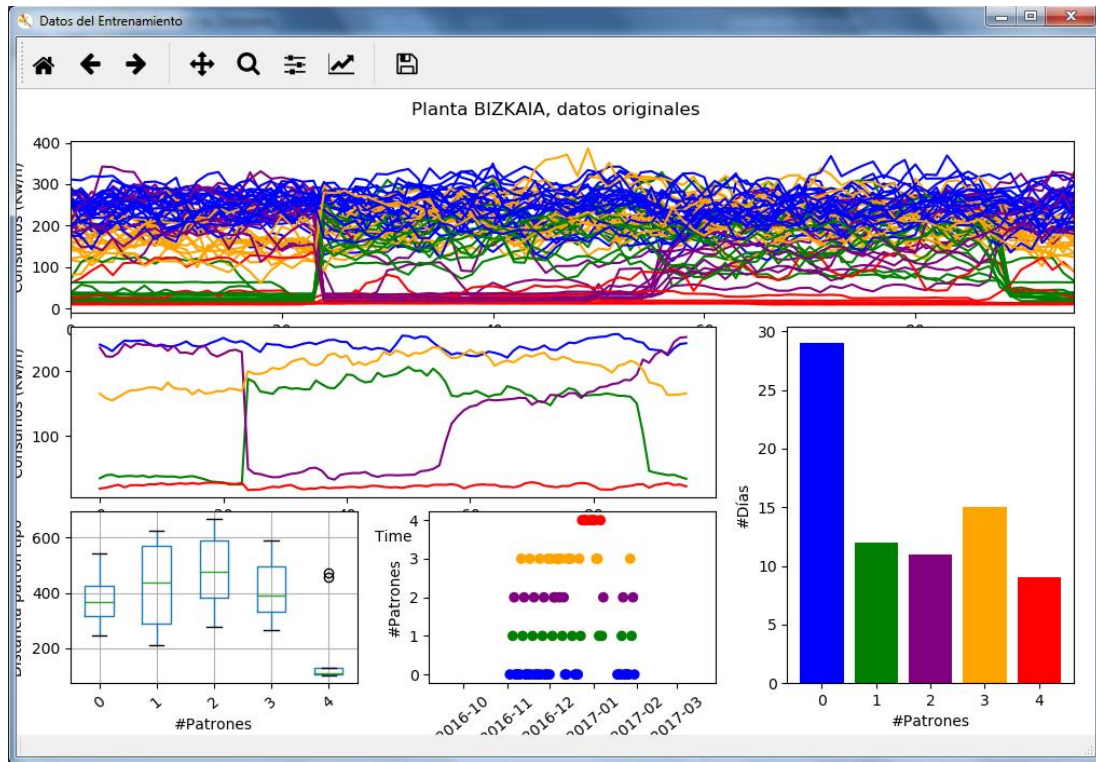


Ilustración 34: Procesamiento de los datos de entrenamiento

9.5. Descripción del algoritmo desarrollado

Por otro lado, se ha decidido que el alumno del presente TFM, con el fin de verificar el correcto funcionamiento de uno de los algoritmos desarrollados por la empresa Tecnia y para seguir avanzando en su aprendizaje en analítica de datos, desarrolle un algoritmo similar al TZAM (Demostrador #1: Zonas multitérmicas de un edificio terciario) pero haciendo uso de las herramientas y módulos disponibles en las librerías SciPy y Sklearn de Python.



Ilustración 35: Mapa del edificio 700. Usado para la toma de los datos del Algoritmo I

Para el desarrollo del algoritmo, será necesaria la validación de los datos disponibles. El algoritmo que se va a desarrollar se ejecutará con los datos disponibles de una de las zonas del mismo edificio 700 de Tecnalia, en Derio, Bizkaia. Como se muestra en la Ilustración 35, la zona 1 dispone a su vez de 12 sub-zonas que se comportan de forma diferente en función de la temperatura y la época del año. Estas sub-zonas podrán ser denominadas también como zonas térmicas.

El algoritmo desarrollado realizará agrupaciones de varias zonas térmicas, en función de su variación de la temperatura con la época del año y del tipo de día. De esta forma, será posible tratar de forma similar a las sub-zonas que tengan un comportamiento similar. Sabiendo cual son las zonas que tienen variaciones similares con la temperatura y conociendo los rangos de temperatura de cada zona, se podrá, controlando las agrupaciones realizadas, asegurar el confort de toda la planta.

Se va a desarrollar un algoritmo que, una vez indicado el periodo del año que se desee analizar (invierno, primavera, verano u otoño) y el tipo de día (entre semana, viernes o fin de semana), represente gráficamente las temperaturas de todas las sub-zonas de la planta de cada día. Atendiendo al número de horas diarias que el sistema esté encendido, se clasificarán los días en: días de jornada normal (sistema encendido 8 horas), días de jornada intensiva (sistema encendido 6 horas) y los fines de semana o festivos (sistema apagado todo el día). A continuación, se dividirá cada día en tres periodos, teniendo en cuenta hora de encendido y apagado del sistema de climatización (horas de antes de encender el sistema, horas mientras que el sistema esta encendido, horas de después de apagar el sistema).

A continuación, se dividirá cada día en tres periodos, en función de la hora de encendido y apagado del sistema de climatización (horas de antes de encender el sistema, horas mientras que el sistema esta encendido, horas de después de apagar el sistema) y se realizará la clusterización de cada periodo de forma independiente. Como resultado se observa que existen tres tipos días que se corresponden a: los días de jornada normal (con el sistema encendido 8 horas), los días de jornada intensiva (con el sistema encendido 6 horas) y los fines de semana o festivos (cuando el sistema está apagado).

Una vez seleccionado el tipo de día y el trimestre del año que se desea analizar se ha procedido a la selección de la técnica de clusterización. Para el algoritmo a desarrollar, será imprescindible que el método de clusterización sea capaz de agrupar las diferentes zonas en función de la variación de la temperatura respecto del tiempo. Aunque existen un gran número de técnicas, se ha

comparado los resultados obtenidos al trabajar con tres técnicas diferentes: k-Means de la librería Sklearn, Agglomerative Clustering de la librería SKlear y fcluster de la librería SPcy. Se han escogido estas tres alternativas debido a que son técnicas que realizan las agrupaciones en función de las distancias existente entre los puntos de estudio.

9.5.1. Librería SciPy

Algunas técnicas de clusterización realizan agrupaciones en base a un número definido de clústeres. Para ello, en primer lugar, será necesario calcular para cada día cual es el número óptimo de clústeres a realizar. Una mala elección de los mismo podrá dar lugar a agrupaciones erróneas de las zonas térmicas. Se han estudiado los diferentes métodos que existen para calcular el número apropiado de clústeres y se ha optado por comparar los métodos del codo [25] y el de silhouette [26]. También se ha hecho uso de los dendrograma para representar las agrupaciones realizadas. Un dendrograma es un tipo de representación gráfica en forma de árbol que organiza y agrupa los datos en subcategorías según su similitud (Ilustración 36).

El método del codo utiliza los valores de la inercia obtenidos tras aplicar el K-means a diferente número de clúster (desde 1 a N clúster), siendo la inercia la suma de las distancias al cuadrado de cada objeto del clúster a su centroide. Al representar en una gráfica los valores obtenidos respecto del número de clúster, el punto donde se aprecia un cambio brusco en la evolución de la inercia será el número óptimo de clúster para el dataset. Por otro lado, el método silhouette se basa en el cálculo de un coeficiente que corresponde a la separación que existe entre los clústeres. Este coeficiente corresponde al promedio del valor silhouette de cada observación y mide el grado de confianza en la asignación de clúster de una observación en particular. Para un número de clústeres k, el ancho de la silueta vendrá dado por el promedio del coeficiente de todas las observaciones.

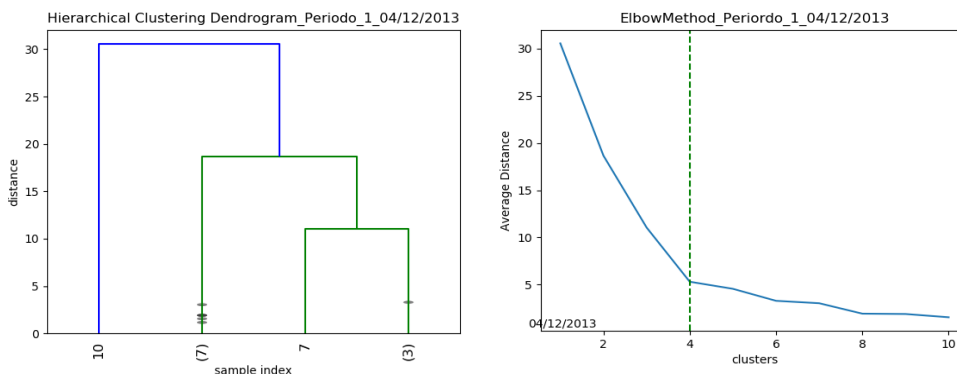


Ilustración 36: Resultado para el Dendrograma y Método del codo

Tras comprobar que ambos métodos dan resultados similares, se ha escogido trabajar con el método del codo. En la imagen superior se muestran los resultados obtenidos al aplicar el método para el 4 de diciembre de 2013 para el periodo en el que el sistema de climatización esta encendido. el número óptimo de agrupaciones calculadas es de 4 clústers.

Una vez que se ha seleccionado el método para la obtención del número óptimo de clúster, se aplicarán las técnicas de clusterización seleccionadas para la agrupación de las zonas multitérmicas. El algoritmo fclusters forma clústeres a partir de una agrupación jerárquica definida por una matriz dada. La agrupación puede hacerse en base a los siguientes criterios:

- Inconsistent: Si un nodo del clúster y todos sus descendientes tienen un valor inconsistente menor o igual a un valor dado t , todos los descendientes de sus hojas pertenecerán al mismo clúster.
- Distance: Forma agrupaciones planas para que las observaciones originales en cada grupo plano no tengan una distancia mayor a un valor dado t .
- Maxclust: Encuentra un umbral mínimo r para que la distancia entre dos observaciones originales en el mismo clúster plano no sea mayor que cierto valor y y que no se formen más de t clúster.

Se han comparado los resultados de los criterios *distance* y *maxclust*. Para el criterio *maxclust*, se le pasará el valor de número de clústeres óptimo calculado con el método del codo. Como se muestra en la siguiente imagen (Ilustración 37), una vez calculado que el valor óptimo es de 4 clústeres, el algoritmo realizará 4 clústeres en función de la distancia entre las observaciones.

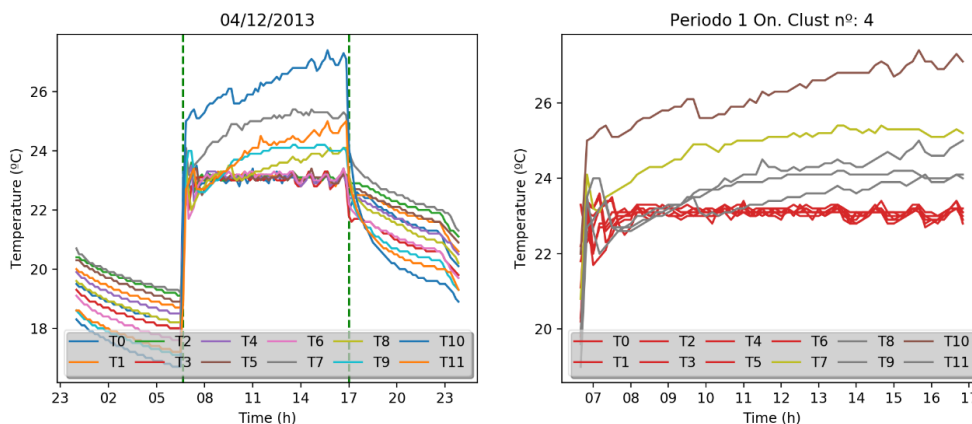


Ilustración 37: Agrupaciones con sistema de climatización encendido con algoritmo fcluster (maxclust)

Por otro lado, el método de fcluster de la librería SciPy también permite hacer la agrupación en base a una distancia introducida como parámetro de entrada. Este valor se establecerá manualmente en base a los resultados obtenidos de las diferentes simulaciones realizadas. Se ha simulado para distancias equivalentes a 0.5°C, 0.8°C, 1.0°C, 1.2°C y 1.8°C. Se agruparán las zonas térmicas que tengan entre si una variación menor o igual a dicho valor. Una vez simulado, se estudiarán los resultados para escoger la magnitud con la que mejor se realice la agrupación de las diferentes zonas.

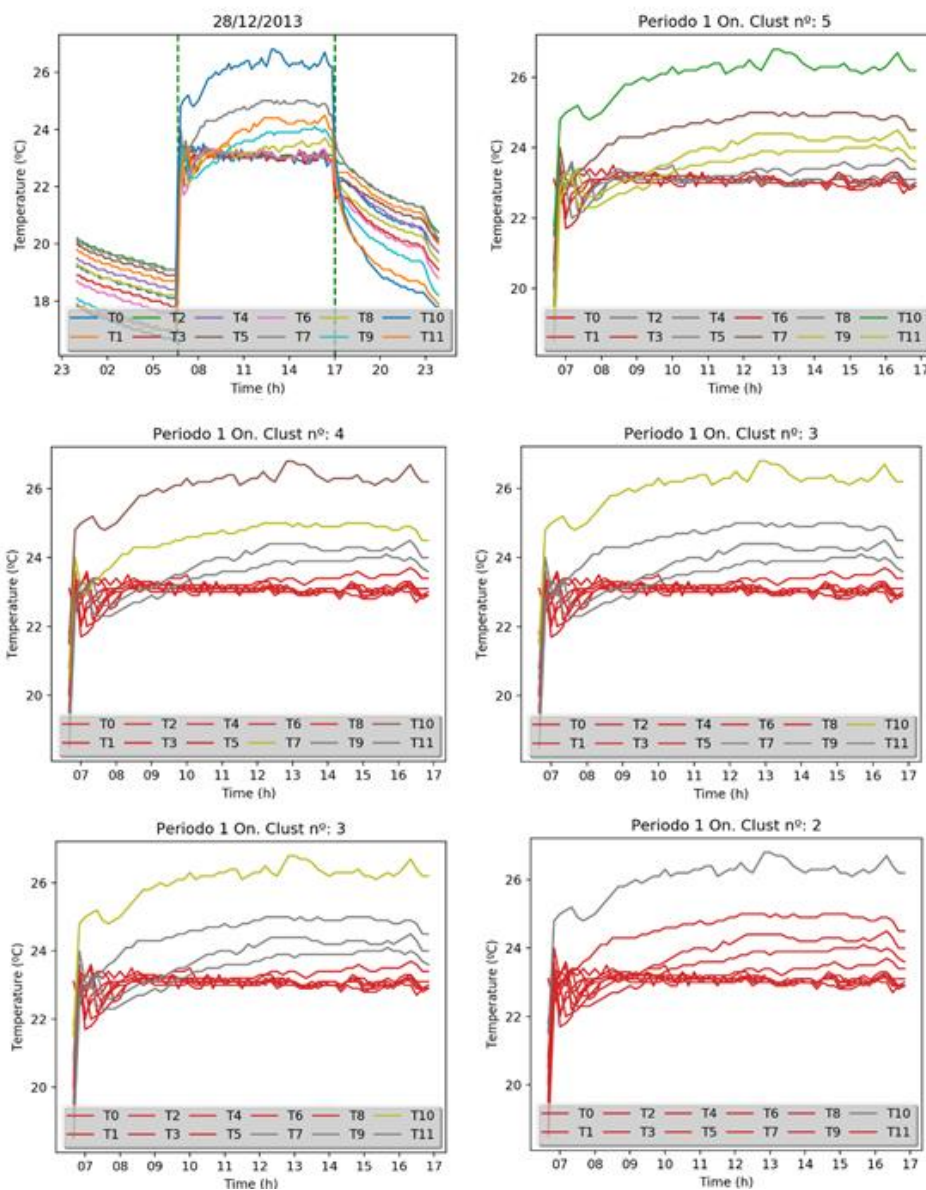


Ilustración 38: Resultados de la clusterización con fcluster de la librería SciPy para distancias de 0.5°C, 0.8°C, 1°C, 1.2°C y 1.8°C

En la primera gráfica de la Ilustración 38 se pueden ver las temperaturas de las 12 zonas térmicas superpuestas para el día escogido. El resto de las gráficas muestran las diferentes agrupaciones realizadas al variar el valor del criterio de agrupación para el periodo horario cuando el sistema HVAC esta encendido.

A medida que disminuye la distancia máxima entre los puntos para considerar las zonas térmicas un mismo clúster, serán necesarios un mayor número de clúster para describir el comportamiento. Al representar los resultados obtenidos, se observa como un índice demasiado alto podrá provocar que se pierda precisión a la hora de agrupar las zonas y podría darse el caso de que de que no se pueda asegurar el confort en todas las zonas de una misma agrupación al mismo tiempo. Por otro lado, un índice demasiado pequeño podría suponer hacer más agrupaciones de las necesarias. Por lo tanto, y en vista a los resultados obtenidos, se ha estimado que una distancia de 1°C es una magnitud suficientemente precisa para realizar las agrupaciones.

Al representar el resto de los periodos de ese mismo día (Ilustración 39) se puede ver cómo, con una distancia de 1°C se consigue agrupar las zonas de diferentes rangos de temperatura con bastante precisión. Por lo tanto, será posible tratar las zonas dentro de un mismo clúster de la misma manera asegurando así el confort en todas ellas al mismo tiempo.

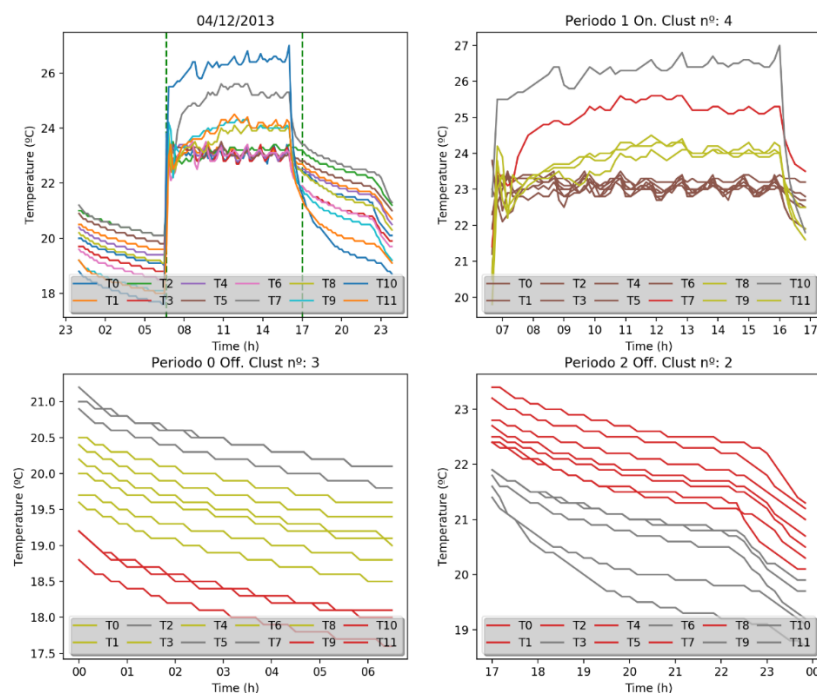


Ilustración 39: Agrupación al aplicar fclusters con criterio de la distancia de 1°C

9.5.2. Librería SkLearn

Las técnicas de K-Means[27] y AgglomerativeClustering[28] de la librería SkLearn calculan los centros de los clústeres y predicen el índice de clúster para cada muestra calculada. Por lo tanto, no será necesario aplicar el método del codo para calcular por anticipado el número óptimo de clústeres.

La técnica K-Means realiza una partición de un conjunto de n observaciones en k grupos en el que cada observación pertenece al grupo cuyo valor medio es más cercano. Por otro lado, la técnica de AgglomerativeClustering fusiona recursivamente el par de clústeres que aumenta mínimamente una distancia dada de enlace. Se han aplicado estos algoritmos y se han comparado los resultados obtenidos para los mismos días que en los ejemplos anteriores.

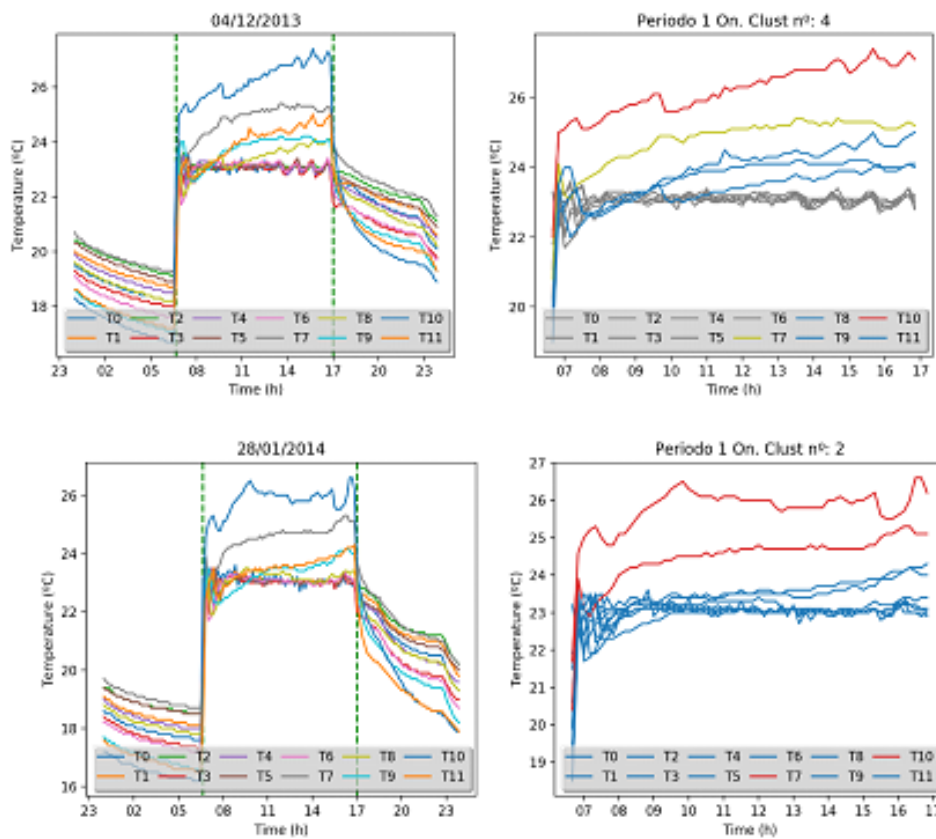


Ilustración 40: Clusterización de método k-Means de la librería SkLearn

Al comparar los resultados obtenidos al aplicar ambos métodos, Ilustración 40 e Ilustración 41 , en ambos casos se obtiene el mismo resultado. Por otro lado, aunque los resultados sean similares a los del caso anterior, donde se trabaja con la librería Scipy con el criterio *distance*,

estos métodos permitirán realizar una herramienta más sistemática que no dependerá tanto de valores introducidos por el usuario. Se tendrá menos margen para establecer la precisión con la que deseé trabajar, pero se asegurará que las agrupaciones están hechas en base a los valores de los datos de los que se dispone y no de estimaciones realizadas por el usuario.

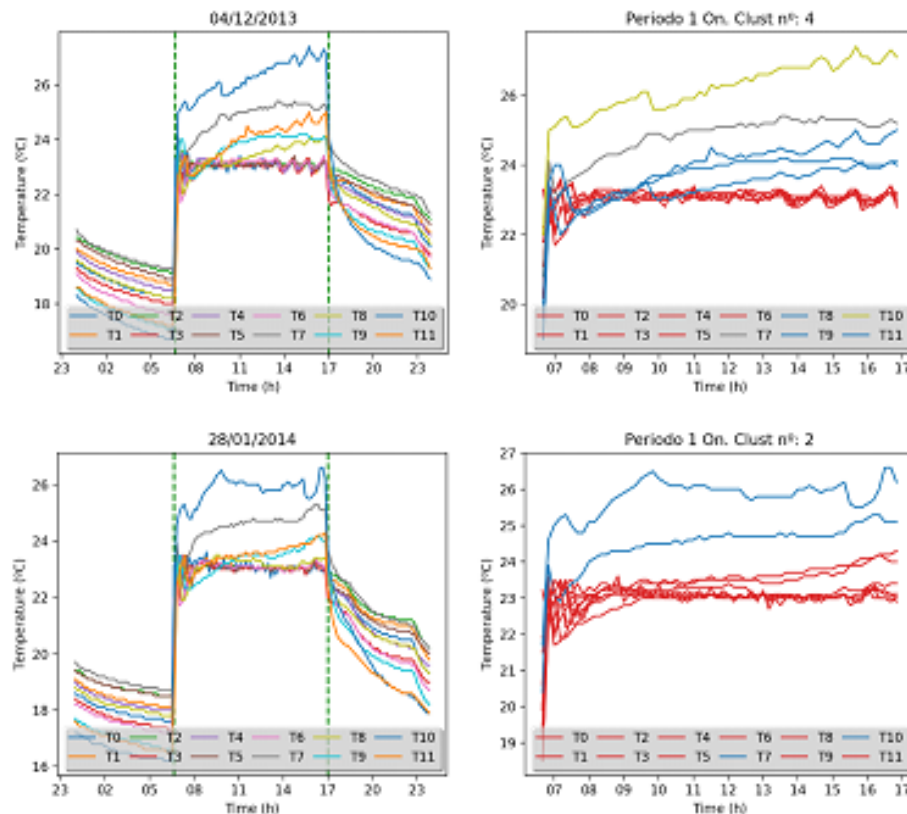


Ilustración 41: Clusterización de métodos AgglomerativeClustering

9.5.3. Correlación

Una vez que procesados todos los días y teniendo un clúster asignado para cada zona térmica, se calculará la correlación de los diferentes periodos (sistema apagado antes de encender, sistema encendido, sistema apagado después de encender) para conocer qué relación existe entre el comportamiento de las diferentes zonas térmicas. Conocer esta relación permitirá tratar las diferentes zonas de forma similar, en cuanto a la climatización se refiere.

La matriz correlación se constituye por los coeficientes de correlación de cada sub-zona con respecto del resto de las sub-zonas. El coeficiente de correlación indicará, por lo tanto, la relación que existe entre el comportamiento de las diferentes sub-zonas.

Para calcular la matriz de correlación se tomará como entrada una matriz con 12 columnas, una por cada sub-zona, y de tantas filas como días hayan sido procesados anteriormente. Dicha matriz tendrá guardado en cada posición el clúster al que perteneció la zona para ese día. De esta forma, gracias al coeficiente de correlación, conoceremos la relación que existe entre el comportamiento de cada zona respecto de las otras.

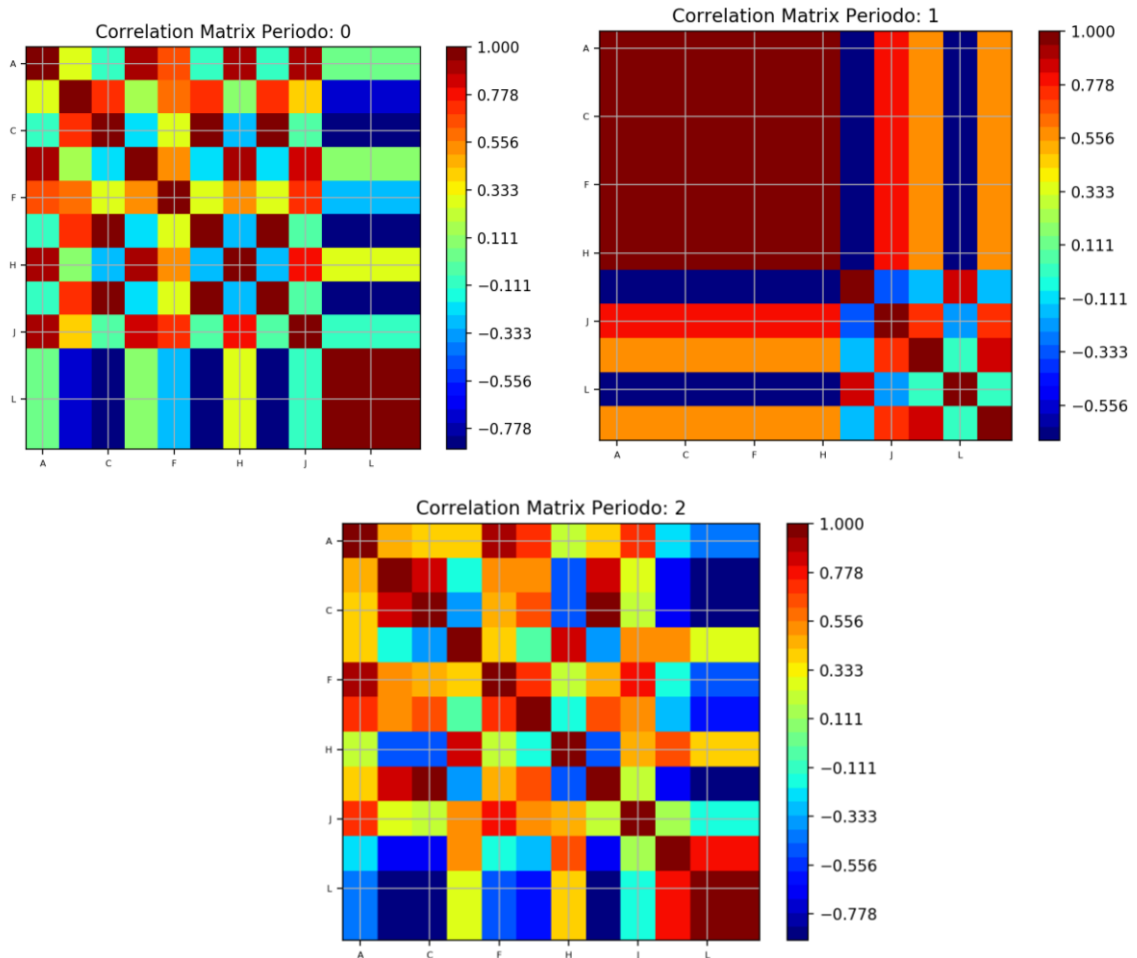


Ilustración 42: Matrices de correlación de los tres periodos para método de fclusters con criterio MaxClust

Cuanto más rojizo sea el color de la celda de la matriz, mayor será la relación existirá entre las zonas térmicas relacionadas. Es por eso por lo que la matriz correlación tiene su diagonal de color marrón, ya que representa a correlación de una zona respecto de ella misma (correlación = 1). Como se puede ver en las diferentes gráficas, la matriz correlación varía en función de si el sistema de climatización este encendido o apagado. También varía en función de la técnica de clusterización aplicada, por ser diferentes el número de agrupaciones calculadas para cada día.

En la siguiente imagen se representa cuáles son las zonas con temperaturas más altas y cuáles las zonas con temperaturas más bajas para el trimestre de invierno. En la gráfica de correlación que muestra los valores para cuando el sistema de climatización esta encendido, se observa como las zonas térmicas que muestran menor relación con respecto de las demás son la zona I y la L. Estas dos zonas son las zonas más caliente y más fría respectivamente, por lo tanto, resulta lógico que muestre un comportamiento menos parecido al del resto de las subzonas.

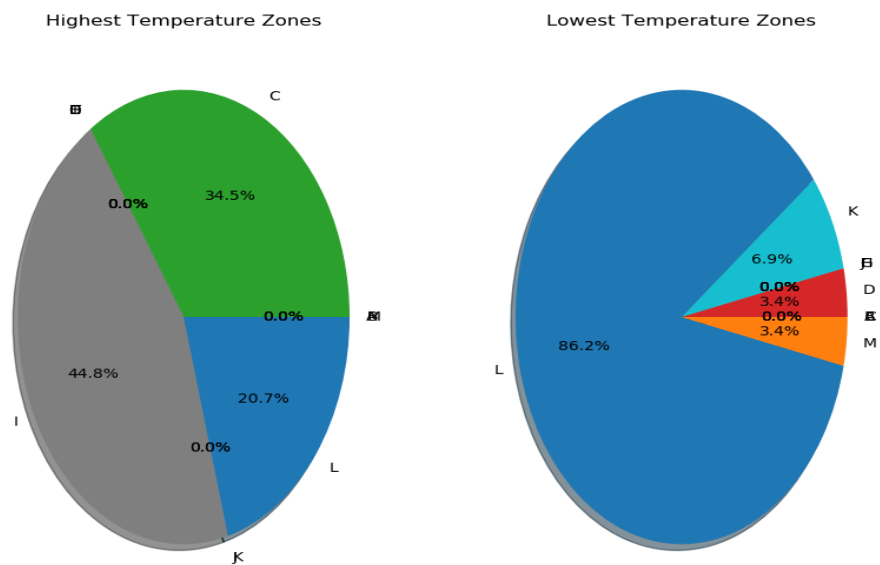


Ilustración 43: Zonas con mayor y menos temperatura diaria (%)

10. Aspectos Económicos

A continuación, se presenta el presupuesto donde se detalla los costes de la realización y puesta en marcha del proyecto.

Para poder calcular el presupuesto total de este Trabajo Fin de Máster se han calculado por separado las horas internas de los trabajadores, los recursos ofimáticos y los gastos requeridos.

10.1. Horas internas

En el desarrollo del TFM han participado dos ingenieros seniors especializados en el área de eficiencia energética en la empresa Tecnalía y un alumno de Máster de Ingeniería Técnica de Telecomunicación con contrato en prácticas en la empresa Tecnalía. Para ello, teniendo en cuenta la duración del proyecto anteriormente planificada, se ha calculado la partida de los trabajadores.

Horas Internas				
	ID	€/hora	Horas	Total
Ingeniero Superior & Director de TFM	IS1	90	148	13.320
Ingeniero Superior	IS2	90	172	15.480
Estudiante	IJ	33	784 ²	12.210
Total				41.010

Tabla 13: Horas Internas

En la siguiente tabla se muestran los costes derivados de los recursos humanos para cada fase de trabajo del proyecto:

² Aunque se ha estimado que será necesario que el alumno dedique un total de 784 horas para la realización del TFM, solo dedicará 370 horas de la jornada de trabajo a la realización del mismo.

Fases	Horas (h)			Importe (€)		
	IS1	IS2	IJ1	IS1	IS2	IJ1
F.1: Curso de formación	30	80	170	2700	7200	5610
F.2: Documentación	30	-	40	2700	-	1320
F.3: Estudio y Selección	-	29	54	-	2610	1782
F.4: Diseño y Desarrollo de la Herramienta	55	55	425	4950	4950	14025
F.5: Simulaciones	25	-	55	2250	-	1815
F.6: Redacción del Informe	8	8	40	720	720	1320

Tabla 14: Fase de los recursos humanos para cada fase de trabajo

10.2. Recursos materiales

Se han calculado, por un lado, los gastos de los materiales activos fijos de la empresa que se podrán utilizar para otros proyectos y, por otro lado los gastos de los materiales que no podrán ser utilizados posteriormente

10.2.1. Amortizables

Para calcular el precio de los materiales que serán amortizables, se ha calculado por un lado la duración total de proyecto y por otro, la vida útil de los materiales utilizados:

DURACIÓN		
Duración Proyecto	8	meses
Duración horas	990	horas

Tabla 15: Cálculo de la duración del TFG

VIDA ÚTIL			
Ordenador	5 años	43800	horas
Licencias (Office365)	1 año	8760	horas
Impresora	10 años	87600	horas
Documentación	20 años	175200	horas

Tabla 16: Cálculo de la vida útil de los materiales

AMORTIZACIONES					
	Unidades	Precio (€) compra/unidad	Horas de uso	€/unidad	Total
Ordenadores	3	1200	650	17,80	53,42
Impresoras	1	1100	30	0,38	0,11
Licencias	3	300	320	10,95	32,88
Documentación		600	75	0,257	0,26
TOTAL					86,67

Tabla 17: Cálculo de las amortizaciones

10.2.2. Gastos

En la partida de gastos, se ha tenido en cuenta los elementos necesarios para llevar a cabo el trabajo que no serán reutilizables:

Total (€)	
	€
Material de oficina	600
Gastos Extra	650
TOTAL	1250

Tabla 18: Cálculo de los gastos

10.2.3. Subcontrataciones

Para este Trabajo Fin de Máster no ha sido necesaria subcontratar a ninguna empresa ni trabajador adicional.

10.3. Resumen

Una vez que se han calculado todas las partidas por separado, se ha aplicado un porcentaje para los costes indirectos del 2%.

Para finalizar, se presenta una tabla resumen en la que se muestran los presupuestos relativos a recursos humanos y recursos materiales, así como el total del presupuesto necesario para la elaboración del proyecto.

PRESUPUESTO TOTAL	
	€
Horas internas	41.010,00
Amortizables	86,67
Gastos	1.250,00
Subtotal	42.346,67
Costes Indirectos (2%)	846,93
TOTAL	43.193,60

Tabla 19: Cálculo del presupuesto total

El presupuesto suma un total de cuarenta y tres mil, ciento noventa y tres con sesenta euros (43.193,60€). Como se puede ver, la mayor parte del presupuesto estará destinada a los sueldos de los trabajadores que van a llevar a cabo este proyecto.

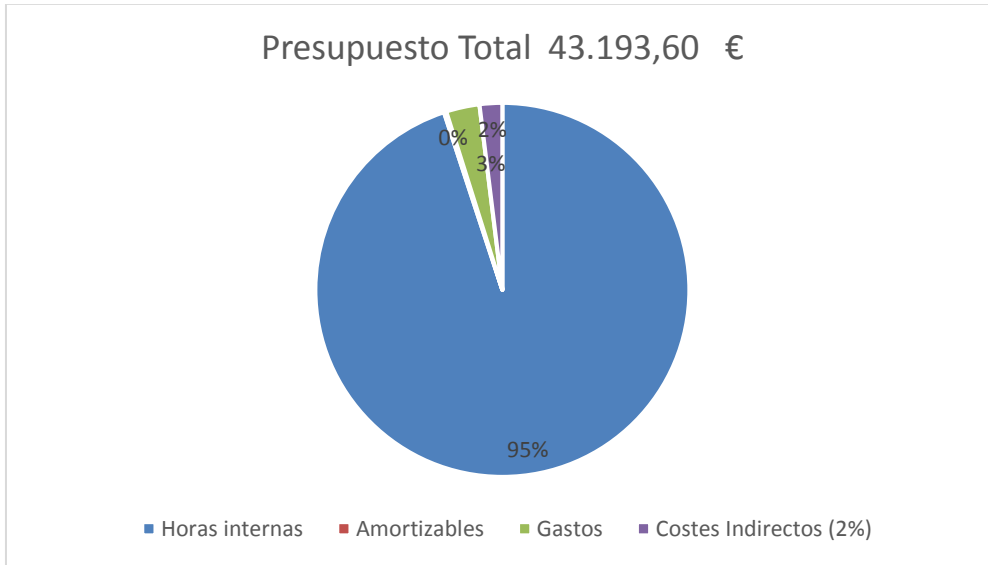


Ilustración 44: Presupuesto final

11. Análisis de Riesgos

En todo proyecto pueden darse una serie de imprevistos que pueden afectar al desarrollo natural del mismo. Por ello, antes de poner en marcha el proyecto, es importante identificar los posibles riesgos que puedan surgir y establecer un plan de contingencia para cada uno de ellos. En esta sección se presentan los riesgos que fueron identificados antes de comenzar con el desarrollo del proyecto.

Se tendrán en cuenta dos factores para clasificar los riesgos identificados:

- Probabilidad de ocurrencia: Probabilidad de que ocurra el imprevisto. Se evalúa en base a tres niveles: baja, media y alta.
- Impacto: Hace Efecto que causaría dicho imprevisto sobre los objetivos del proyecto en caso de que ocurriese. Se evalúa en base a tres niveles: bajo, medio y alto.

11.1. Retrasos

El incumplimiento de los plazos establecidos para el desarrollo de las distintas tareas. Esto supondrá un retraso en la fecha de finalización del proyecto y un incremento en los costes. Debido a que el trabajo realizado no solo depende del trabajo del alumno que realiza el proyecto y que los trabajadores de Tecnalía tienen una alta carga de trabajo, es un riesgo que estará presente en todo proyecto.

- Probabilidad: Alta
- Impacto: Medio

Debido a la alta probabilidad de ocurrencia de este riesgo, antes de planificar el proyecto se decidió planificar con margen suficiente como para poder hacer frente a pequeñas demoras que pudiesen surgir sin causar un gran impacto en el desarrollo del mismo.

11.2. Bajas de participantes del proyecto

Otro riesgo que está presente en cualquier proyecto es la posibilidad de que alguno de los trabajadores tenga que ausentarse debido a alguna enfermedad o motivos personales. El impacto

podría ser alto si la ausencia del trabajador se alargase demasiado, por ejemplo, unas semanas.

- Probabilidad: Baja
- Impacto: Medio

Si la duración de la baja es corta, no tendría un gran impacto sobre el proyecto, puesto que únicamente retrasaría su finalización un poco más de lo previsto. Una baja de larga duración es necesario un plan de contingencia que permita tener la situación bajo control. Gracias a que Tecnalía dispone de más trabajadores con conocimientos suficientes como para ayudar a llevar a cabo los diferentes proyectos, cualquiera de ellos podría supervisar el proyecto durante la ausencia de uno de los responsables del proyecto.

11.3. Costes imprevistos

Un riesgo muy común en la elaboración de cualquier proyecto son los gastos no presupuestados que producen un desajuste en el presupuesto previsto inicialmente. En este proyecto en concreto la probabilidad de tener gastos no previstos es bastante baja. La ejecución de este proyecto no supone un gran impacto económico y la probabilidad de tener gastos no previstos con anterioridad es bastante baja. Las pequeñas desviaciones que pudiesen surgir debido a gastos imprevistos no serían de gran relevancia.

- Probabilidad: Baja
- Impacto: Bajo

Ante una situación como esta, el plan de contingencia consiste en solicitar financiación externa que permita a la división de Tecnalía disponer de más recursos económicos, lo que permitiría destinar cierta parte a la finalización de este proyecto.

11.4. Desarrollo de un Smart Energy Lab difícil de manejar

Al desarrollar el Smart Energy Lab puede darse el caso de que la herramienta realizada sea demasiado compleja para aquellos trabajadores de Tecnalía que no hayan trabajado previamente con entornos de programación o herramientas similares. El Notebook de Jupyter es una aplicación poco común entre los trabajadores, por lo que será necesario que los trabajadores que vayan a

hacer uso de la herramienta dispongan de algún tipo de ayuda para aprender a manejar la misma y así sean capaces de ejecutarla de forma autónoma.

- Probabilidad: Media
- Impacto: Alta

Para evitar este problema se realizará un manual de usuario y un video tutorial donde se explicarán los pasos a seguir y de forma detalla la forma correcta de ejecución de los demostradores.

11.5. Matriz probabilidad/ Impacto

Después de analizar las distintas amenazas, se incluyeron todas ellas en una matriz para poder analizar de un modo simple y visual la magnitud probabilidad/impacto de cada uno de los riesgos identificados.

		IMPACTO		
		BAJA	MEDIA	ALTA
PROBABILIDAD	BAJA	Costes imprevistos	Bajas de participantes del proyecto	
	MEDIA			
	ALTA		Retrasos Desarrollo de un Smart Energy Lab difícil de manejar	

Ilustración 45: Matriz Probabilidad-Impacto

12. Conclusiones

En este documento se ha presentado el trabajo llevado a cabo para el diseño del Smart Energy Lab, una herramienta gráfica para la integración de cuatro algoritmos desarrollados en la División de Energía y Medio Ambiente de la empresa Tecnalía, que aglutina una serie de aplicaciones para el subsector energético, Uso Final de Energía. Además, incluye un nuevo algoritmo para el sector de la eficiencia energética que hace uso de técnicas *machine learning* y analítica de datos, aportando por tanto un valor añadido adicional al complementar los cuatro algoritmos previamente existentes.

Con el fin de desarrollar una solución interactiva y visual, fácil de usar y entender para el usuario final, se ha optado por trabajar con los Notebooks de Jupyter. De esta forma, se ha diseñado una herramienta de ayuda en la que los trabajadores podrán apoyarse para explicar y desarrollar conceptos y objetivos de los diferentes proyectos desarrollados en la empresa, permitiendo así que los potenciales clientes del sector energético entiendan las ventajas de las tecnologías digital propuestas.

Por otro lado, el hecho de desarrollar la herramienta ha dado al alumno la oportunidad de comprender los conceptos más importantes que se trabajan en DEMA y de adaptar y realizar pequeñas mejoras en los algoritmos aplicados. Además, el haber desarrollado su propio algoritmo con aplicación en el sector energético, le ha servido para aprender y trabajar con tecnologías de utilidad más general como son *machine learning* y analítica de datos.

Se han escogido los cuatro proyectos que mejor engloban los diferentes aspectos que se trabajan en la división y se ha adaptado la herramienta gráfica según las características de cada proyecto. Esta solución se ha llevado a cabo mediante la plataforma Jupyter y lenguaje de programación Python. Se ha creado una única base de datos donde estarán almacenados los datos recopilados por los diferentes sensores e integrado los diferentes algoritmos que se van a utilizar. La herramienta permitirá variar los parámetros de entrada más relevantes de cada proyecto para luego visualizar los resultados obtenidos y calcular, mediante algoritmos de analítica de datos, soluciones de ahorro energético.

Con vistas al futuro, la herramienta realizada podrá extenderse y usarse como referencia para seguir adaptando el resto de algoritmos ya desarrollados en la división. En el análisis de alternativas, debido al tiempo estipulado para la realización del TFM, de los 6 proyectos planteados se escogieron 4 para adaptar y desarrollar este trabajo. Gracias al trabajo realizado, será más sencillo adaptar y desarrollar nuevos Notebooks para el resto de los proyectos.

Además, con el propósito de ayudar a la persona que haga uso de los Notebooks, se ha incluido en un anexo el manual de usuario donde se explicará de forma detallada el funcionamiento de la herramienta.

13. Bibliografía

- [1]. TecNALIA. Inspiring Business, [<https://www.tecnalia.com/es/>]. [Consulta: 20-02-2018]
- [2]. "El hombre, responsable del cambio climático" | Edición impresa | EL PAÍS, [https://elpais.com/diario/2007/02/03/sociedad/1170457201_850215.html]. [Consulta: 19-07-2018]
- [3]. Acción por el clima - EUROPA | Unión Europea, [https://europa.eu/european-union/topics/climate-action_es]. [Consulta: 18-05-2018]
- [4]. Comisión Europea, Acción por el clima, [https://ec.europa.eu/clima/index_es]. [Consulta: 03-03-2018]
- [5]. Plataforma Tecnológica Española de Eficiencia Energética, [<http://www.ptee.org/es/>]. [Consulta: 20-03-2018]
- [6]. La importancia de las normas ISO en las organizaciones, [<https://www.isotools.org/>]. [Consulta: 03-03-2018]
- [7]. La industria 4.0 en la mejora de la eficiencia energética - Conciencia Eco, [<https://www.concienciaeco.com/2017/09/11/la-industria-4-0-en-la-mejora-de-la-eficiencia-energetica/>]. [Consulta: 09-03-2018]
- [8]. ISO 50001 Gestión Energética | BSI Group, [<https://www.bsigroup.com/es-ES/ISO-50001-Gestion-Energetica/>]. [Consulta: 19-06-2018]
- [9]. ¿Por qué implantar un sistema de gestión de energía? | Zemsania Global Group, [<https://www.zemsania.com/sistema-de-gestion-de-energia/>]. [Consulta: 14-05-2018]
- [10]. B&R: Sistemas de control de procesos, [<https://www.br-automation.com/es/productos/sistemas-de-control-de-procesos/>]. [Consulta: 10-02-2018]

- [11]. Geothermal HVAC Heating and Cooling | Cair Heating & Cooling | Louisville KY, [<http://cairhvac.com/geothermal-heat/>]. [Consulta: 01-03-2018]
- [12]. Bombas de Calor | Web Oficial Fujitsu, [<http://www.disfrutael Fujitsu.com/bombas-de-calor/>]. [Consulta: 04-05-2018]
- [13]. Bombas de calor, tipos y aplicaciones - Novelec Blog, [<http://blog.gruponovelec.com/climatizacion/bombas-de-calor-tipos-y-aplicaciones/>]. [Consulta: 01-03-2018]
- [14]. ¿Qué es Big data y para que sirve? Una introducción a Big data – Bit, [<https://www.bit.es/knowledge-center/que-es-big-data-introduccion-a-big-data/>]. [Consulta: 04-05-2018]
- [15]. Las 7 V del Big data: Características más importantes – IIC, [<http://www.iic.uam.es/innovacion/big-data-caracteristicas-mas-importantes-7-v/#volumen>]. [Consulta: 04-05-2018]
- [16]. Capitulo I. Sistemas y Modelados, [<ftp://ftp.unicauca.edu.co>]. [Consulta: 09-01-2018]
- [17]. Hastie, T., Tibshirani, R., and Friedman, J. (2008). The Elements of Statistical Learning. Data Mining, Inference and Prediction. Springer. Chapter 2 and Chapter 14. <https://web.stanford.edu/~hastie/Papers/ESLII.pdf>
- [18]. Conceptos básicos de Machine Learning -, [<https://cleverdata.io/conceptos-basicos-machine-learning/>]. [Consulta: 09-01-2018]
- [19]. Aprendizaje Inductivo: Árboles de Decisión - Fernando Sancho Caparrini, [<http://www.cs.us.es/~fsancho/?e=104>]. [Consulta: 09-01-2018]

- [20]. Project Jupyter | Home, [<http://jupyter.org/>]. [Consulta: 15-02-2018]
- [21]. MathWorks - Makers of MATLAB and Simulink, [https://es.mathworks.com/?s_tid=gn_logo].
[Consulta: 15-02-2018]
- [22]. Scikit-learn: machine learning in Python — scikit-learn 0.19.1 documentation, [<http://scikit-learn.org/stable/>]. [Consulta: 16-05-2018]
- [23]. SciPy — SciPy v1.1.0 Reference Guide, [
<https://docs.scipy.org/doc/scipy/reference/index.html>] [Consulta: 16-05-2018]
- [24]. Information Security information, news and tips, [<https://searchsecurity.techtarget.com/>].
[Consulta: 22-05-2018]
- [25]. Selección del número óptimo de Clusters – Jarroba, [<https://jarroba.com/seleccion-del-numero-optimo-clusters/>]. [Consulta: 18-05-2018]
- [26]. informe_final_introi_ni_lena.pdf, [
https://eva.fing.edu.uy/file.php/514/ARCHIVO/2011/TrabajosFinales2011/informe_final_introi_ni_lena.pdf]. [Consulta: 18-05-2018]
- [27]. Sklearn.cluster.KMeans — scikit-learn 0.19.1 documentation, [<http://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html>]. [Consulta: 21-05-2018]
- [28]. Sklearn.cluster.AgglomerativeClustering — scikit-learn 0.19.1 documentation, [<http://scikit-learn.org/stable/modules/generated/sklearn.cluster.AgglomerativeClustering.html#sklearn.cluster.AgglomerativeClustering>]. [Consulta: 21-05-2018]

14. Anexos

A continuación, se ha anexoado una imagen donde se visualiza el diagrama de bloques de la base de datos diseñada para guardar los datos recopilados de los diferentes sensores.

Por otro lado, se ha anexoado el Manual de Usuario del trabajo, donde se explica de manera detallada cómo hacer a la Máquina Virtual y como hacer uso de cada una de las interfaces desarrolladas para cada práctica.

ANEXO I

Base de Datos

A continuación, se ha anexo un diagrama de bloques del diseño de la base de datos donde estarán recopilados los datos de los dispositivos y sensores de los que se tomarán los datos. Estos datos serán lo que se introducirán como parámetros de entrada a los diferentes demostradores que se han desarrollado.

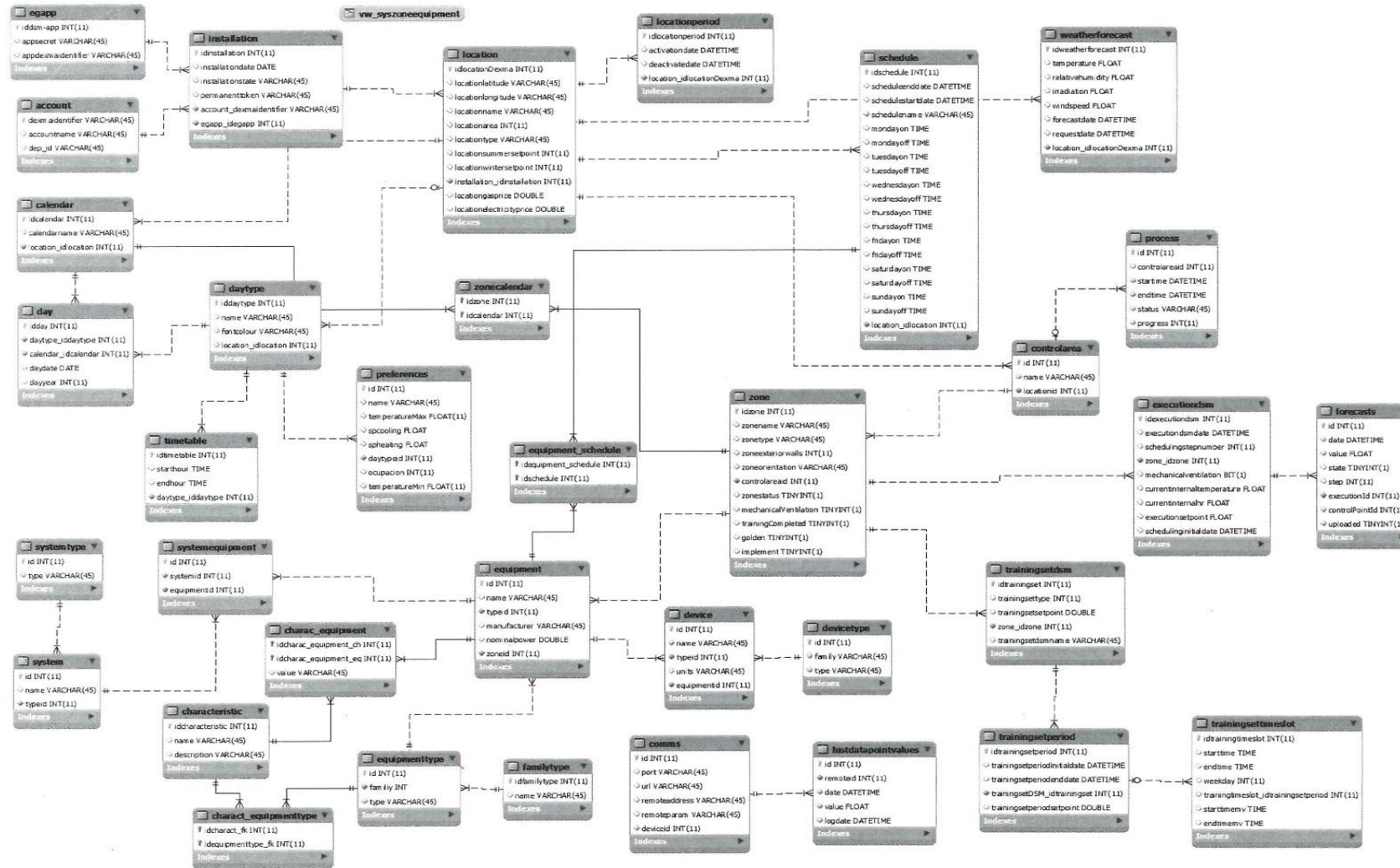


Ilustración 46: Esquema de Bloques de la Base de datos

ANEXO II

Manual de Usuario

Lanzar Jupyter

Como ya se ha explicado en el presente documento, la aplicación que se ha desarrollado estará disponible en una máquina virtual de la DMZ de Tecnalia. Para poder acceder a el menú de la aplicación, el usuario, una vez conectado a la red de la empresa (vía wifi, físicamente o través de la VPN), deberá acceder a la siguiente URL:

http://172.26.252.195:8890/tree/SMART_ENERGY_LAB

En la siguiente imagen se muestra el menú al que se accede mediante la URL dada.

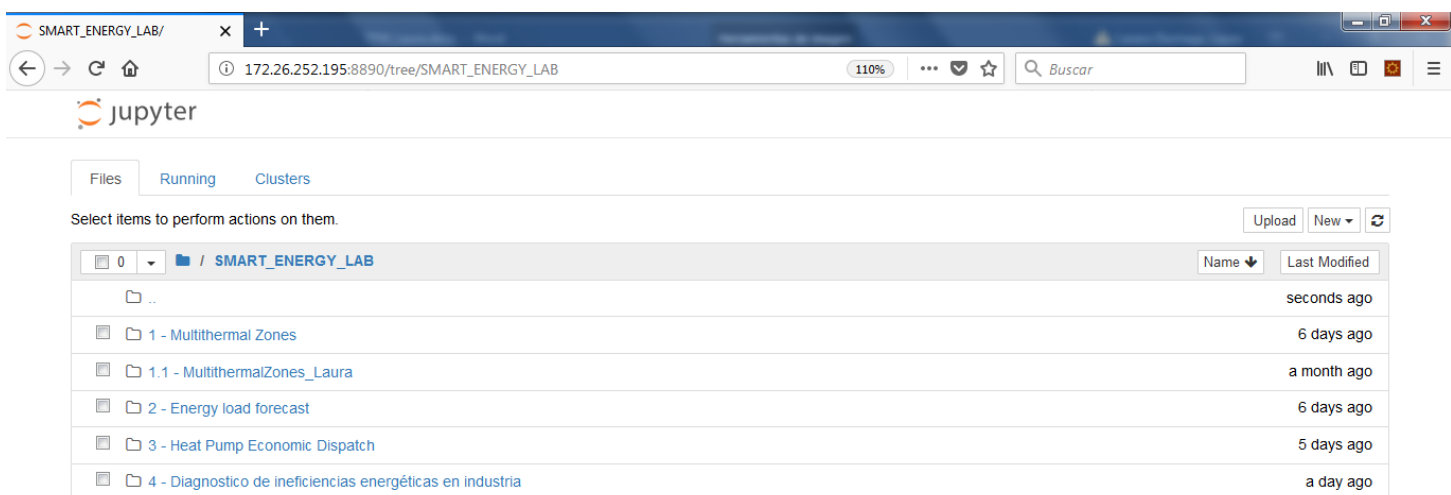


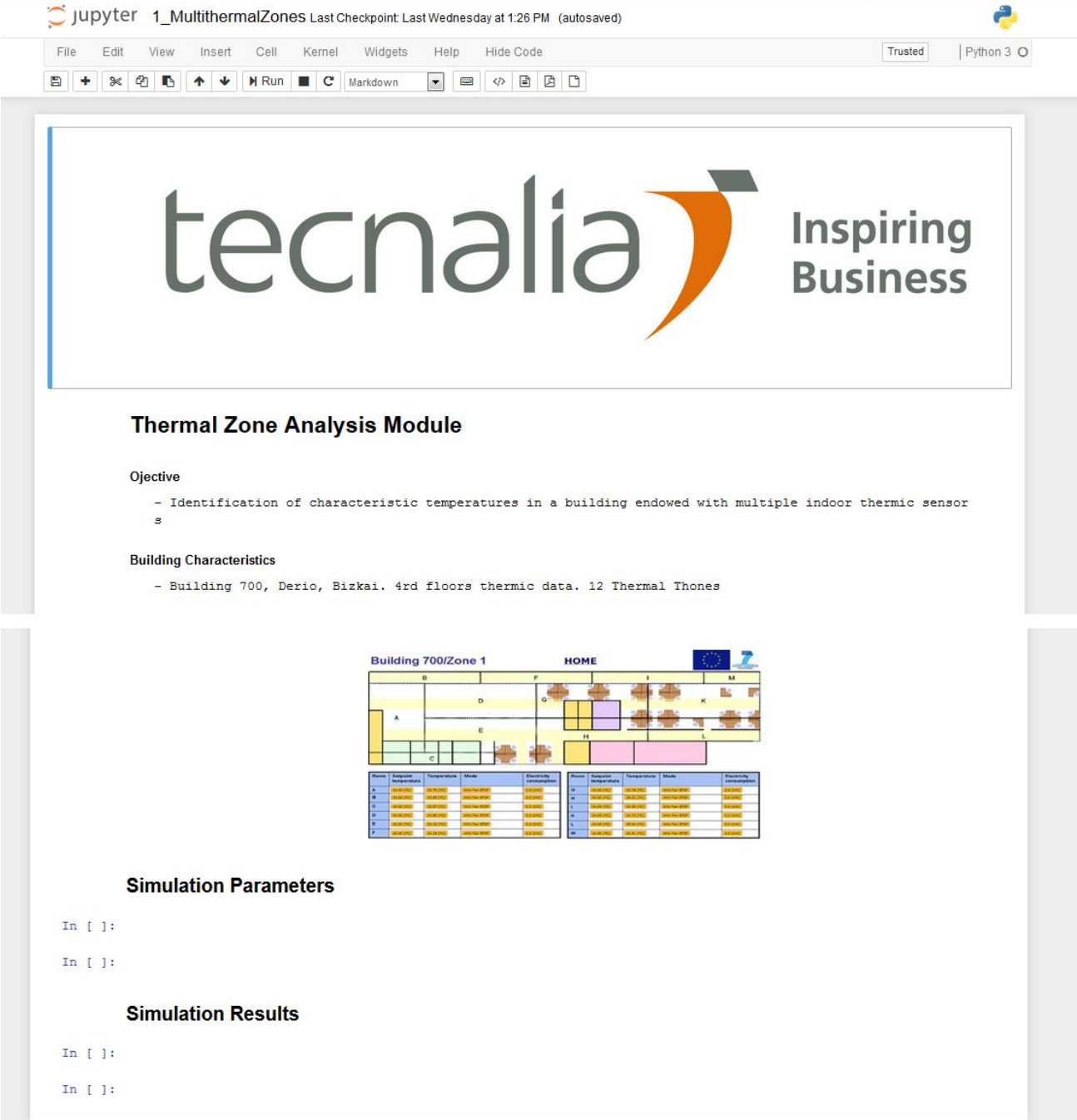
Ilustración 47: Menú de la Herramienta diseñada

Aunque las máquinas virtuales de la DMZ se reinicien periódicamente, nos aseguramos de que la aplicación esté disponible para ser ejecutada en todo momento.

Una vez en el menú, el usuario podrá acceder al escenario que desee clicando en la carpeta correspondiente. A continuación se accederá al Notebook de Jupyter, el cual tendrá el mismo nombre que la carpeta pero con extensión *.ipynb*. Con el fin de facilitar el aprendizaje y el manejo de la herramienta, se ha seguido un mismo patrón para el diseño de los diferentes escenarios desarrollados.

Ejecución de los Escenarios

A continuación, se describirá la ejecución de uno de los escenarios desarrollados:



jupyter 1_MultithermalZones Last Checkpoint: Last Wednesday at 1:26 PM (autosaved)

File Edit View Insert Cell Kernel Widgets Help Hide Code Trusted Python 3

tecnalia Inspiring Business

Thermal Zone Analysis Module

Objective

- Identification of characteristic temperatures in a building endowed with multiple indoor thermic sensors

Building Characteristics

- Building 700, Derio, Bizkai. 4rd floors thermic data. 12 Thermal Zones

Building 700/Zone 1 HOME

Room	Surface Temperature	Temperature	Mode	Electricity Consumption	Room	Surface Temperature	Temperature	Mode	Electricity Consumption
A	20.000000	20.000000	Indoor	0.000000	H	20.000000	20.000000	Indoor	0.000000
B	20.000000	20.000000	Indoor	0.000000	I	20.000000	20.000000	Indoor	0.000000
C	20.000000	20.000000	Indoor	0.000000	J	20.000000	20.000000	Indoor	0.000000
D	20.000000	20.000000	Indoor	0.000000	K	20.000000	20.000000	Indoor	0.000000
E	20.000000	20.000000	Indoor	0.000000	L	20.000000	20.000000	Indoor	0.000000
F	20.000000	20.000000	Indoor	0.000000	M	20.000000	20.000000	Indoor	0.000000
G	20.000000	20.000000	Indoor	0.000000					

Simulation Parameters

In []:

In []:

Simulation Results

In []:

In []:

Ilustración 48: Notebook del Escenario 1. Thermal Zone Analysis Module

Como se muestra en la Ilustración 48, cada una de las aplicaciones dispondrá de una breve introducción al caso práctico y a continuación el código a ejecutar. Por razones de diseño se ha decidido esconder el código de forma que solo queden visibles los parámetros de entrada y salida que sean relevantes para los exponentes y oyentes.

Las casillas que deban de ser ejecutables tendrán una indicación de **In []** a su izquierda. La ejecución se realizará de arriba abajo, clicando sobre la casilla a ejecutar (de forma que la casilla quede seleccionada con un recuadro azul. Ver Ilustración 49) y seleccionando el botón de **Run**.

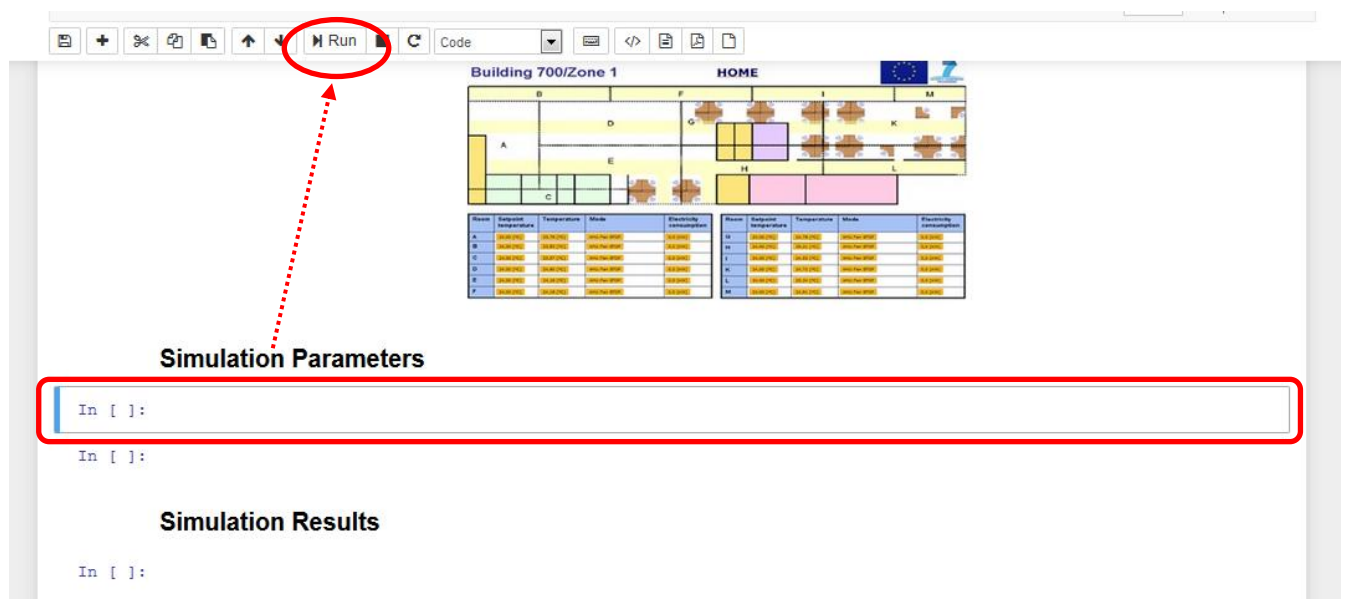


Ilustración 49: Pasos para la ejecución del Notebook

- In[*] -> Indica que la celda está ejecutándose.
- In[1] -> Un número entre los corchetes indicará que la ejecución ha terminado y el usuario podrá seguir ejecutando la próxima celda.

Las celdas servirán o bien para pedir los parámetros de simulación al usuario, o bien para representar gráficamente los resultados. Las celdas que sirvan para pedir la entrada de datos vendrán con valores por defecto que podrán ser fácilmente modificados simplemente clicando o variando los parámetros correspondientes:

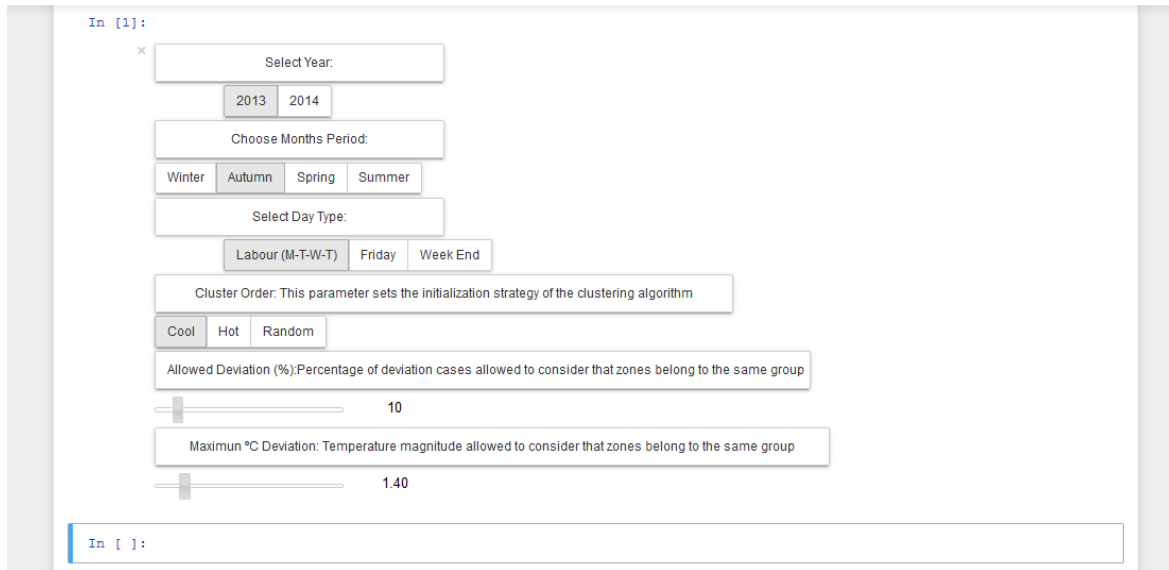


Ilustración 50: Ejecución del Notebook. Introducción de parámetros para la simulación

***Atención:** Una vez modificados los parámetros de entrada, habrá que seguir ejecutando las casillas posteriores. El volver a pulsar el botón de run para la casilla donde se han modificado los parámetros de entrada, hará que se vuelvan a establecerse los parámetros por defecto.

Los resultados obtenidos serán representados gráficamente y podrán ser guardados siempre que el usuario lo desee. Además, la herramienta dará la opción de realizar zoom a la imagen para estudiar en detalle los resultados obtenidos en la simulación.

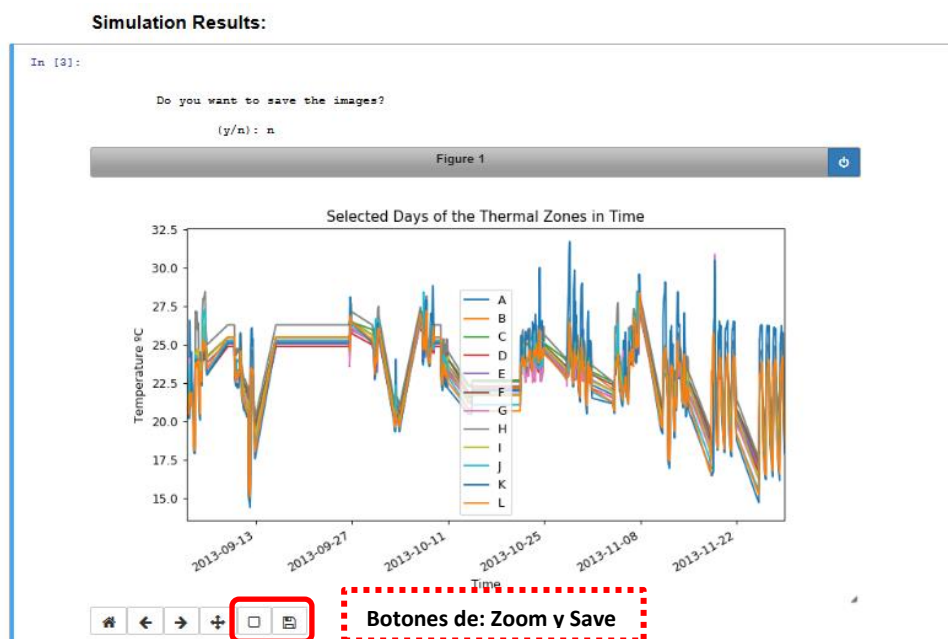


Ilustración 51: Ejecución del Notebook. Visualización de los resultados

Reinicialización y Exit

- Reiniciar Notebook: Menú superior **Kernel-> Restart & Clean Output**

En caso de tener problemas con la herramienta o de que por alguna razón el tiempo de ejecución sea más elevado de lo normal, se recomienda reiniciar el Kernel. De esta forma se finalizará la ejecución actual y se iniciará una nueva.

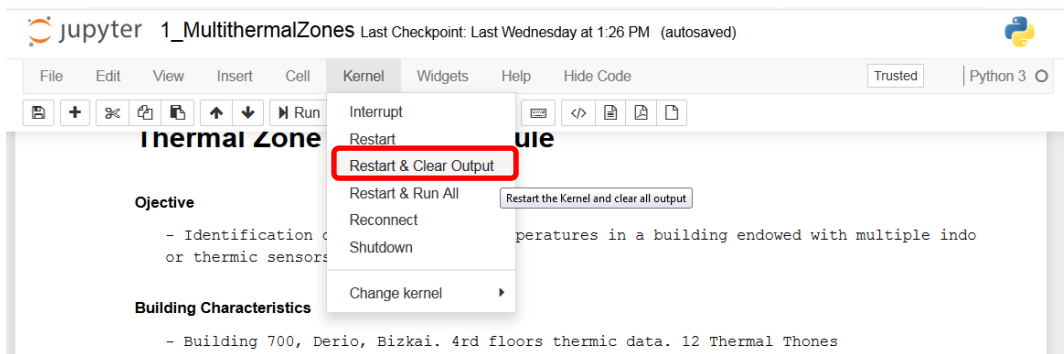


Ilustración 52: Reinicio del Kernel

- Cerrar Notebook: Menú superior: **Kernel-> Restart & Clean Output**. Menú superior: **File -> Save & CheckPoint**. Menú superior: **File -> Close & Halt**

Una vez finalizada la ejecución, es conveniente apagar el Kernel que se haya utilizado, limpiar el Notebook y guardarlo. De esta forma, el próximo usuario que vaya a ejecutar el escenario dispondrá de la herramienta sin ningún valor o resultados de las simulaciones anteriores.

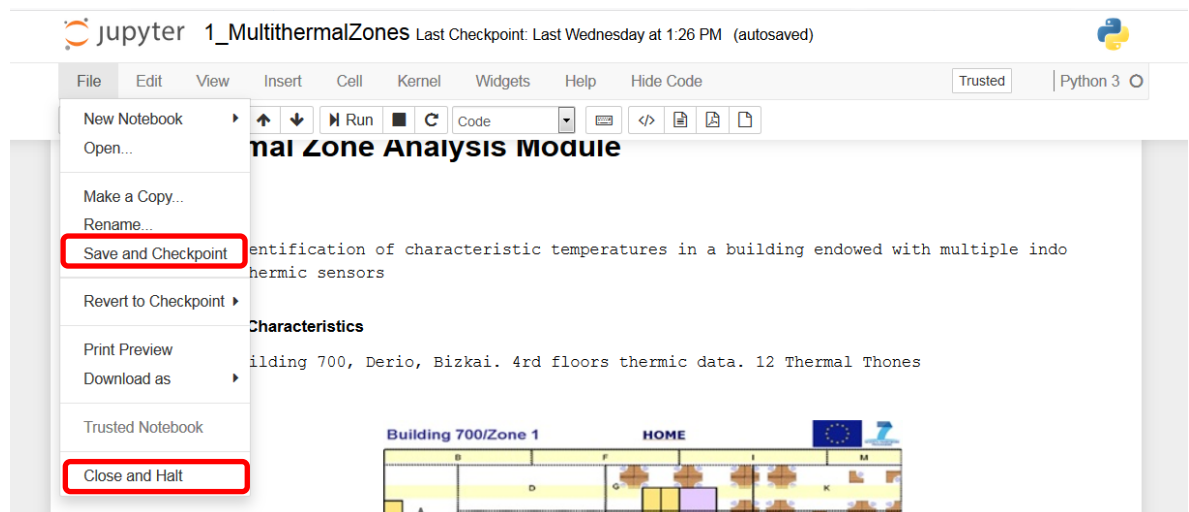


Ilustración 53: Fin de la ejecución & Cierre de la aplicación

