



Universidade de Vigo

Trabajo Fin de Máster

---

# Detector de fallos en maquinaria mediante Técnicas Estadísticas y Machine Learning

---

Manuel Domínguez Basteiro

Máster en Técnicas Estadísticas

Curso 2021-2022



## Propuesta de Trabajo Fin de Máster

<p><b>Título en galego:</b> Detector de fallos en maquinaria mediante Técnicas Estadísticas e Machine Learning</p>
<p><b>Título en español:</b> Detector de fallos en maquinaria mediante Técnicas Estadísticas y Machine Learning</p>
<p><b>English title:</b> Machinery fault detector using Statistical Techniques and Machine Learning</p>
<p><b>Modalidad:</b> Modalidad B</p>
<p><b>Autor/a:</b> Manuel Domínguez Basteiro, Universidade da Coruña</p>
<p><b>Director/a:</b> Salvador Naya Fernández, Universidade da Coruña; Javier Tarrío Saavedra, Universidad da Coruña</p>
<p><b>Tutor/a:</b> Jose González Areas, Cofrico S.L</p>
<p><b>Breve resumen del trabajo:</b></p> <p>Desarrollo de un software para la detección de fallos en distintos tipos de maquinaria mediante el manejo del Big Data, el uso de técnicas de Machine Learning y de Control Estadístico de la Calidad; el cual permite accionar procedimientos automáticos que solventen dichas deficiencias o prevengan fallos futuros.</p>
<p><b>Recomendaciones:</b></p>
<p><b>Otras observaciones:</b> A petición de COFRICO por la protección de los datos de sus clientes y dado que el software desarrollado en este TFM se esta implementando como un desarrollo de I+D de la empresa, se tomó la decisión de no publicar el trabajo original. En su defecto, se proporciona el siguiente resumen del proyecto original.</p>



# Resumen del trabajo completo

## Resumen en español

Este Trabajo de Fin de Máster tiene por objeto la detección y análisis las anomalías en el contexto de la refrigeración, además de proponer soluciones estadísticas al problema del control de la calidad de las instalaciones y maquinaria de la empresa COFRICO. Para ello se analizaron los datos de la empresa y se propusieron modelos de gráficos de control multivariantes. Estas soluciones fueron implementadas en un software específico que por requerimientos de la propia empresa está realizado en Python.

Dicho software permite detectar fallos en el funcionamiento de distintos tipos de maquinaria y activar diferentes medidas automáticas para solventarlos. Además, el software sienta las bases para una futura predicción de este tipo de fallos mediante el estudio de las series temporales.

El programa funciona mediante el manejo de grandes volúmenes de datos (Big Data), el uso de técnicas de Machine Learning (ML) y el uso de diferentes técnicas estadísticas del Control de Procesos y del Control Estadístico de la Calidad (CEC).

Se empieza haciendo una presentación de COFRICO y de las motivaciones y necesidades para la creación de este programa y su implantación en la plataforma GRADHOC, de monitorización y gestión preventiva de instalaciones frigoríficas. Luego se hace una presentación de GRADHOC y se comenta el funcionamiento de una instalación frigorífica.

Después se habla sobre el desarrollo del programa de detección de atípicos, desde la idea inicial al desarrollo final. Se continúa explicando la estadística utilizada y las técnicas de Machine Learning y de Control Estadístico de la Calidad utilizadas en este Trabajo de Fin de Máster.

En el siguiente capítulo, se comentan las herramientas, programas y librerías usadas y como fue la implantación del programa en Python.

Finalmente, se comentan y discuten los resultados obtenidos durante la implantación del software en diferentes instalaciones reales, las líneas futuras de desarrollo y se muestran las principales funciones de código.

## English abstract

This Master's Thesis aims to analyze anomalies in the context of refrigeration and to suggest statistical solutions to the quality control problem of these anomalies. For this purpose, company data were analyzed and multivariate control chart models were proposed. These solutions were implemented in a specific software developed in Python due to the company's own requirements.

This software allows to detect malfunctions in the operation of different types of machinery and to activate different automatic measures to solve them. In addition, this software lays the foundations for a future prediction of this type of failure through the study of time series.

The program works through the management of large volumes of data (Big Data), the use of Machine Learning (ML) techniques and the use of different statistical techniques of Process Control and Statistical Quality Control (SQC).

The first step is an introduction of COFRICO and its motivations and needs in order to create this program and its implementation in the GRADHOC platform for monitoring and preventive ma-

nagement of refrigeration facilities. Then, a detailed description of GRADHOC and the operation of a refrigeration facility is made.

Afterwards, the development of the outlier detection program is discussed from the initial idea to its final development. During this chapter, the statistics, the Machine Learning and Statistical Quality Control techniques used in this Master's Thesis are explained.

The next chapter discusses the tools, programs and packages used and how the program was implemented in Python.

Eventually, the results obtained during the implementation of this software in different real facilities, the future lines of development and the main code functions are shown.

# Introducción

Se comentan ahora parte de los capítulos del trabajo completo, para hacer un resumen del mismo. Se han eliminado las partes que explican la lógica del software y las técnicas usadas a petición expresa de la empresa.

También se ha eliminado el apéndice final, que contenía las principales funciones de Python del programa.

## Presentación de la empresa

COFRICO es una empresa que nace en Burela, en 1985 con la finalidad de diseñar, proyectar, ejecutar y mantener instalaciones de refrigeración y climatización tanto industriales como comerciales. En los 37 años posteriores a su fundación se ha expandido hasta convertirse en una de las principales empresas del país en este sector. Parte de su éxito reside en su apuesta por innovación, la eficiencia energética y su compromiso con el medio ambiente, siendo pionera en varias innovaciones en este ámbito.

En su afán por las nuevas tecnologías, la empresa se encuentra en desarrollo de un sistema de monitorización, mantenimiento predictivo y automatización de todas sus instalaciones. Esto se realiza mediante una plataforma pionera llamada GRADHOC, la cual integra toda la información en tiempo real de las variables enérgicas y las medidas de toda la maquinaria de cada instalación. Además realiza una optimización y un mantenimiento preventivo y predictivo. Actualmente este sistema está saliendo al mercado.

Los valores de la empresa son la innovación, el compromiso medioambiental, la eficiencia energética y la calidad de servicio. Para más información revisar la página web de COFRICO [?].

## Motivación del problema

Es en esta nueva plataforma GRADHOC, en lo referente al mantenimiento predictivo y a la apuesta por la innovación y las nuevas tecnologías, donde surge la motivación para el problema afrontado en este TFM. Se comenta brevemente en esta sección su planteamiento y en capítulos posteriores profundizaremos tanto en GRADHOC como en el problema.

Una parte importante del cualquier proceso industrial es el mantenimiento de la maquinaria, ya que el deterioro de los componentes puede acabar en averías que paren el funcionamiento y obliguen a realizar tareas de reparación. Esto aumenta los costes de la empresa o del cliente, al tener que realizar más intervenciones y comprar e instalar recambios.

Por tanto, un elemento de gran valor para la plataforma predictiva son los programas de detección y predicción de fallos, para esta labor, COFRICO ha querido desarrollar softwares diferentes. La empresa tenía la idea de realizar dos de ellos desde el punto de vista de la física y haciendo alusión a su apuesta por la innovación se pensó en incorporar a un estadístico para que aportara los conocimientos de esa ciencia.

Teniendo esto en cuenta, COFRICO recurrió al Máster en Técnicas Estadísticas en búsqueda de un alumno en prácticas que apoyara al equipo en sus diferentes áreas, siendo el primero con esta formación del equipo.

Gracias a los conocimientos aprendidos en asignaturas como el Control Estadístico de la Calidad y Aprendizaje Estadístico y basándose en una idea inicial del tutor de la empresa, José González, sobre la posibilidad de usar técnicas de predicción para ver desviaciones en el funcionamiento de la maquinaria, se pensó en el siguiente diseño de un software predictivo:

El software detecta fallos en las máquinas de las instalaciones frigoríficas y para ello se usa la gran cantidad de datos medidos de cada máquina de la instalación en tiempo real y guardados en las bases de datos. Dichos valores reales consisten en medidas de la temperatura, presión y demás magnitudes de los distintos componentes de las máquinas.

Partiendo de la idea de que un modelo de Machine Learning entrenado con un patrón de datos predice mal si se le introduce otro diferente, se pensó en entrenar un modelo con los datos de una máquina en una franja de tiempo en el que su funcionamiento se considera correcto y luego ir añadiendo datos actuales al modelo, para comprobar si los errores en la predicción aumentan. Si el funcionamiento se mantiene correcto, estos no aumentarán, pero si el funcionamiento se degrada, sí.

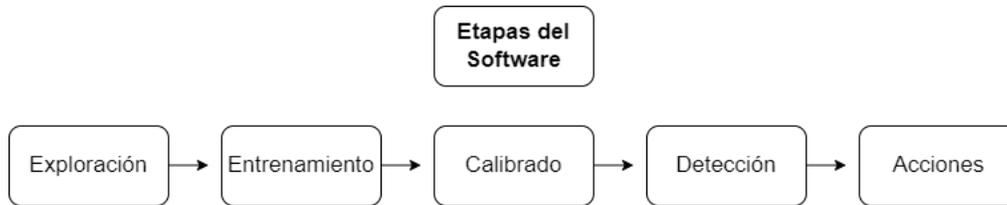


Figura 1: Diagrama de flujo del planteamiento inicial.

El programa consta de 5 fases.

1. Realizando una exploración de las variables más relevantes de la máquina, se seleccionan dos franjas de tiempo donde se considere que la máquina tiene un buen funcionamiento para entrenar y luego calibrar el modelo.
2. Se entrena el modelo de Machine Learning para predecir las variables de mayor importancia.
3. Se calibra el modelo con unos datos diferentes a los del entrenamiento para determinar su buena habilidad de predicción y estimación de los límites que permitan la detectar una posible degradación del funcionamiento.
4. Se fusionan las variables mediante técnicas estadísticas, que dan lugar a gráficos de control que permiten, mediante un procesamiento de los mismos determinar cuando se degrada o cambia el funcionamiento de la máquina.
5. Si la máquina empieza a degradar su funcionamiento, se inician los procesos automáticos para poner en conocimiento de este suceso a los técnicos o a la instalación para realizar un mantenimiento preventivo de la misma.

# Plataforma monitorizadora e instalaciones frigoríficas

En este capítulo, se comenta cómo son las instalaciones y las máquinas con las que actualmente funciona el software propuesto en este TFM, aunque este se puede aplicar a todo tipo de maquinaria. También se hablará de cómo actualmente se hace el control y el tratamiento de los datos, mediante el uso de GRADHOC. Se empezará describiendo superficialmente las funciones de GRADHOC para ver como encaja el software desarrollado en este TFM en ella.

Se incluye el capítulo tal cual aparece en el TFM.

## GRADHOC

COFRICO es una empresa de refrigeración tanto industrial como comercial, por lo tanto GRADHOC gestiona ambos tipos de instalaciones. El software de detección de fallos está pensado para ser utilizado en los dos tipos de instalaciones, sin embargo, dado que aun se encuentra en desarrollo, este solo se ha implementado en instalaciones comerciales. Concretamente, analizaremos su implementación en el control de un supermercado.

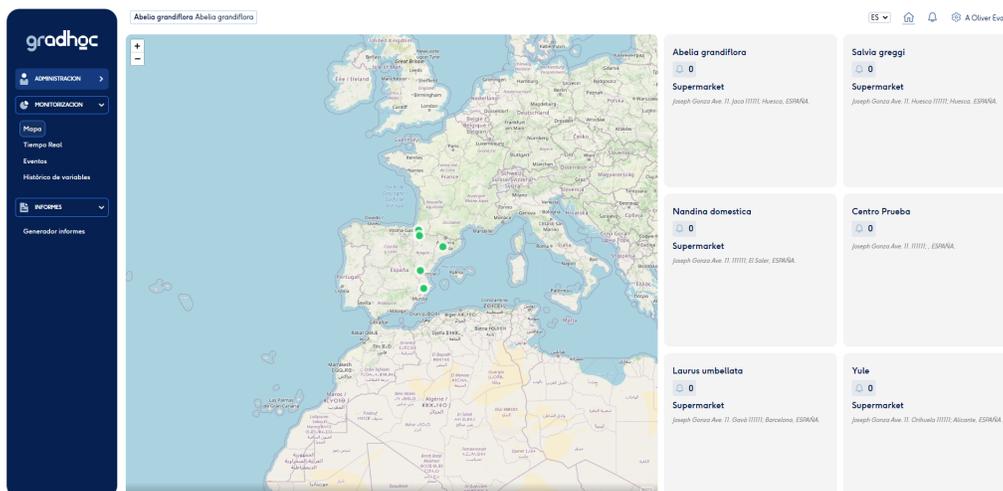
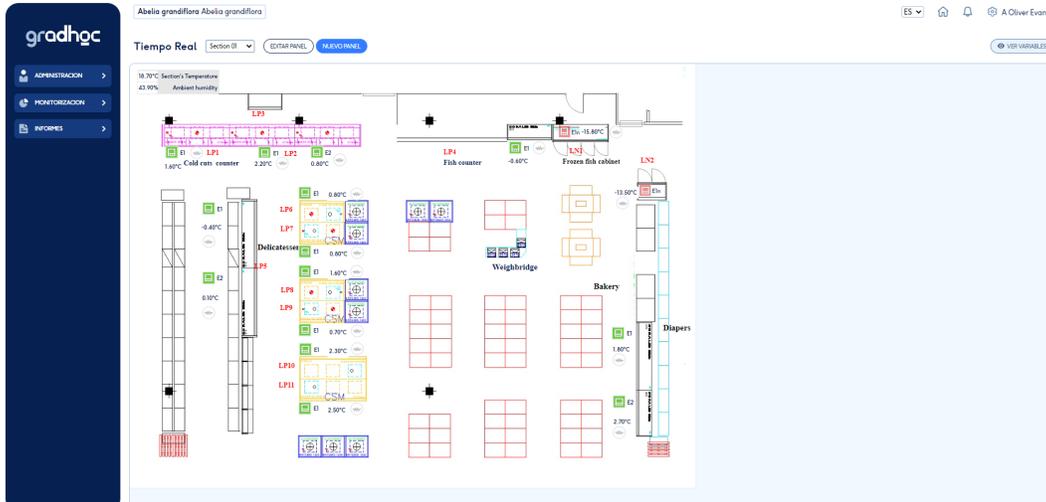


Figura 2: Mapa de las instalaciones.

Se muestra a continuación mediante imágenes extraídas de GRADHOC, un funcionamiento superficial de sus funciones y se verá donde encaja el software de detección de atípicos dentro de la plataforma. Esto servirá para mostrar luego una instalación de ejemplo dónde se ejecuta el programa. Para más información sobre GRADHOC, se puede consultar su página web [?].

GRADHOC se pensó como un aplicación web, a la cual se accede con usuario y contraseña. Una vez dentro, el cliente tiene acceso a toda la información de las instalaciones monitorizadas en tiempo real, los eventos predictivos o preventivos y a un generación de informes a demanda. Nada más acceder el cliente ve un mapa con todas las instalaciones que tiene monitorizadas, como se muestra en la Figura 1.1.

En los valores a tiempo real se pueden visualizar los datos más importantes de la maquinaria mediante diagramas y mapas de la instalación. Se pueden ver tanto un mapa con las vitrinas de un supermercado, como un diagrama de la maquinaria general del frío de la instalación. Además, tiene un graficador integrado que permite ver el histórico de las variables de cada máquina.



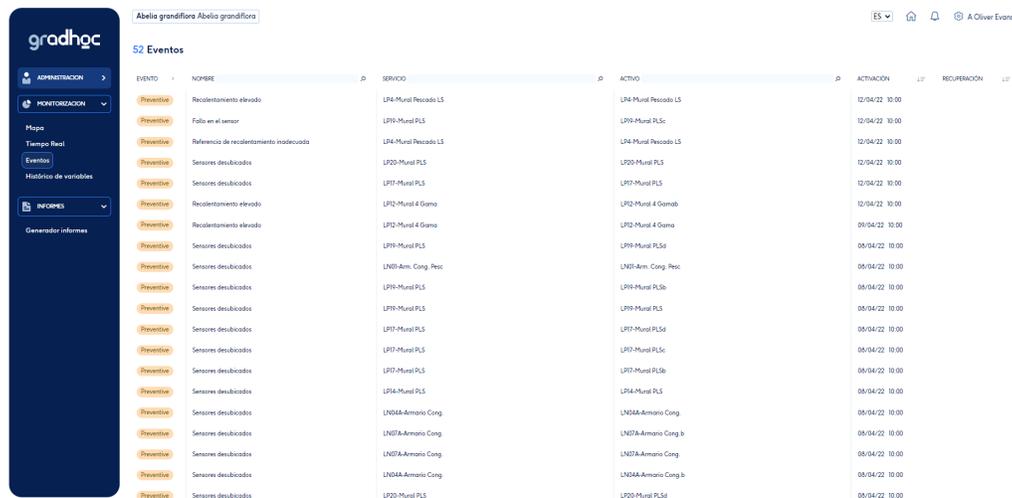
(a) Vitrinas en tiempo real.



(b) Histórico de variables.

Figura 3: Valores en tiempo real de Gradhoc.

Las partes en donde encajará el trabajo realizado en este TFM son el generador de informes y la sección de eventos. En la sección de eventos es donde aparecen los avisos que los distintos tipos de programas predictivos y preventivos detectan, para que el cliente pueda estar avisado y ver los informes asociados si es de su interés. En la sección de informes se puede ejecutar a petición el software de detección de atípicos para una máquina y una franja de tiempo concretas.



EVENTO	NOMBRE	SERVICIO	ACTIVO	ACTIVACION	RECUPERACION
Preventiva	Recalentamiento elevado	LPA-Mural Pezudo LS	LPA-Mural Pezudo LS	12/04/22 10:00	
Preventiva	Fallo en el sensor	LPI0-Mural PLS	LPI0-Mural PLSc	12/04/22 10:00	
Preventiva	Referencia de recalentamiento inadecuada	LPA-Mural Pezudo LS	LPA-Mural Pezudo LS	12/04/22 10:00	
Preventiva	Sensores desactivados	LPI0-Mural PLS	LPI0-Mural PLS	12/04/22 10:00	
Preventiva	Sensores desactivados	LPI0-Mural PLS	LPI0-Mural PLS	12/04/22 10:00	
Preventiva	Recalentamiento elevado	LPI2-Mural 4 Gama	LPI2-Mural 4 Gama b	12/04/22 10:00	
Preventiva	Recalentamiento elevado	LPI2-Mural 4 Gama	LPI2-Mural 4 Gama	09/04/21 10:00	
Preventiva	Sensores desactivados	LPI0-Mural PLS	LPI0-Mural PLSd	08/04/22 10:00	
Preventiva	Sensores desactivados	LN01-Arm. Cong. Resc	LN01-Arm. Cong. Resc	08/04/22 10:00	
Preventiva	Sensores desactivados	LPI0-Mural PLS	LPI0-Mural PLSb	08/04/22 10:00	
Preventiva	Sensores desactivados	LPI0-Mural PLS	LPI0-Mural PLS	08/04/22 10:00	
Preventiva	Sensores desactivados	LPI0-Mural PLS	LPI0-Mural PLSd	08/04/22 10:00	
Preventiva	Sensores desactivados	LPI0-Mural PLS	LPI0-Mural PLSc	08/04/22 10:00	
Preventiva	Sensores desactivados	LPI0-Mural PLS	LPI0-Mural PLSb	08/04/22 10:00	
Preventiva	Sensores desactivados	LPI6-Mural PLS	LPI6-Mural PLS	08/04/22 10:00	
Preventiva	Sensores desactivados	LN04A-Armario Cong	LN04A-Armario Cong	08/04/22 10:00	
Preventiva	Sensores desactivados	LN07A-Armario Cong	LN07A-Armario Cong b	08/04/22 10:00	
Preventiva	Sensores desactivados	LN07A-Armario Cong	LN07A-Armario Cong	08/04/22 10:00	
Preventiva	Sensores desactivados	LN04A-Armario Cong	LN04A-Armario Cong b	08/04/22 10:00	
Preventiva	Sensores desactivados	LPI0-Mural PLS	LPI0-Mural PLSd	08/04/22 10:00	

Figura 4: Tabla de eventos

## Instalaciones frigoríficas

A continuación se explicará como es una instalación comercial y las distintas máquinas que la componen, centrándonos en los evaporadores, que es en la cual se ha empezado a desarrollar el software de detección estadístico. Para la realización de esta sección fue de especial interés el Trabajo de Fin de Máster, José González Areas (2019) [?].

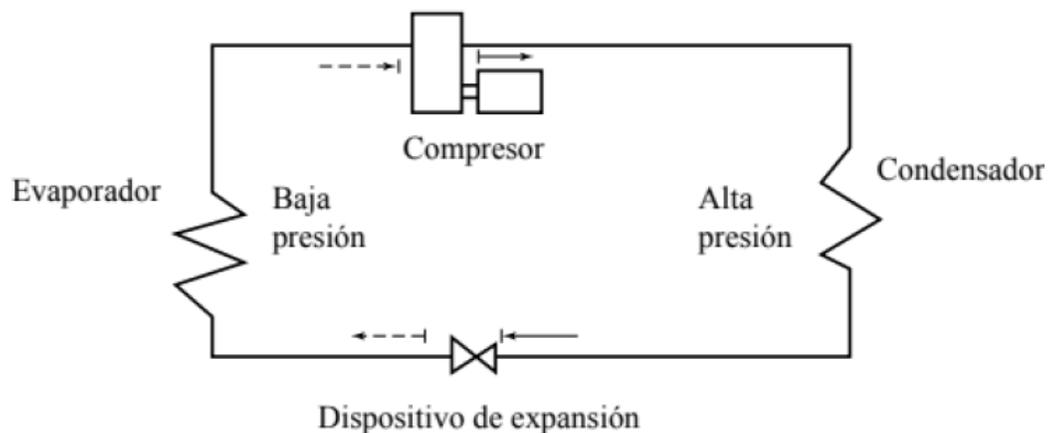


Figura 5: Esquema de una instalación frigorífica.

En una instalación frigorífica básica intervienen los siguientes componentes: compresor, condensador, válvulas de expansión y evaporadores. Normalmente en la sala de máquinas se encuentran los compresores, condensadores y las válvulas que se dedican a enfriar el refrigerante, que luego mediante tubos es enviado a los evaporadores de las vitrinas y muebles de la instalación, un supermercado por ejemplo. Las etapas de un ciclo de refrigeración son las siguientes:

Primero, se debe comprimir el gas refrigerante, esto se hace con el compresor, el cual esta acoplado a un motor eléctrico. Con esto consigues que el gas ocupe un menor volumen y por tanto que aumente su temperatura. Luego, el gas comprimido a alta presión es conducido por el conducto de descarga del

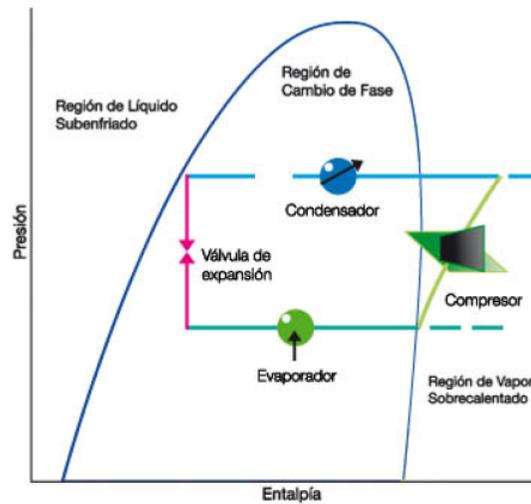


Figura 6: Ciclo del frío.

compresor hacia el condensador.

En esta máquina es donde el gas transfiere su calor al medioambiente y se condensa, pero manteniendo la presión. En cuanto el líquido refrigerante abandona el condensador, este atraviesa un filtro para extraer las impurezas del mismo. A partir de este punto, el proceso depende del refrigerante utilizado, los más usados son el amoníaco ( $\text{NH}_3$ ) y el dióxido de carbono ( $\text{CO}_2$ ).

Para el caso del amoníaco ( $\text{NH}_3$ ), con una válvula de expansión que genera una pequeña caída de presión, se condensa parte del líquido, lo que produce una caída de temperatura en él. Luego esta mezcla de líquido y gas frío es almacenada en un tanque. En la parte superior del tanque está el gas y en la inferior el líquido, ambos en contacto y con una temperatura y presión iguales. Con esto se consigue que el compresor aspire de nuevo el gas de arriba para empezar de nuevo el ciclo y el líquido sea enviado a los evaporadores.



Figura 7: Ejemplo de un evaporador.

Para el dióxido de carbono ( $\text{CO}_2$ ), existe una válvula de expansión que genera una gran caída de la presión, de forma que el líquido pasa a gas: este cambio reduce la temperatura del gas.

La última parte del ciclo del frío sucede cuando el fluido frío llega a los evaporadores. Un evaporador es un conducto cuya superficie es perpendicular a laminas metálicas, como se ve en la Figura 1.6.

Lo que se desea es que el aire, el cual se quiere enfriar, circule por entre esas láminas, entrando en contacto con el conducto del refrigerante. En este proceso el material absorbe una gran cantidad de calor del medio externo (aire de la cámara o el producto de su interior) con lo que se baja la temperatura del medio que se quiere enfriar. Para maximizar el contacto entre el aire y el serpentín, que es el nombre del recorrido del refrigerante, el evaporador puede incluir un ventilador para asegurar el paso del aire. Al terminar, con unos tubos se devuelve el refrigerante al compresor para iniciar el proceso de nuevo.

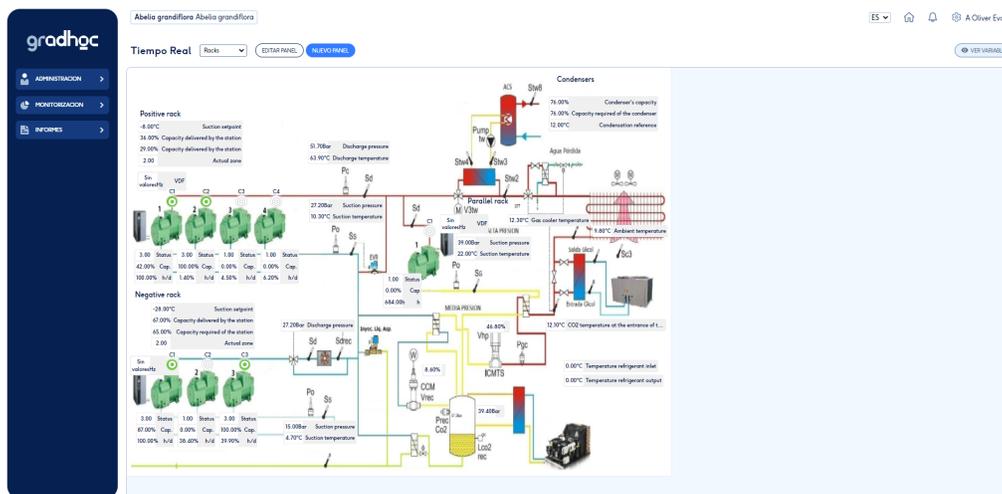


Figura 8: Esquema de una instalación frigorífica comercial.

Finalmente, se muestra en la Figura 1.7 cómo es una instalación frigorífica real de un supermercado que se monitoria con GRADHOC. Aunque cuenta con más compresores, condensadores y válvulas de expansión, el funcionamiento es el mismo que el descrito anteriormente.



# Capítulos del Trabajo de Fin de Máster

En esta parte se explican superficialmente los capítulos que componen el TFM original. Que a petición de COFRICO se ha decidido no mostrarlos públicamente.

## Programa de detección de anomalías

En este capítulo es donde se presentaba el software detector desarrollado en el Trabajo de Fin de Máster. Se comentaba el planteamiento inicial del programa, desde la idea original del tutor de la empresa, hasta la versión final desarrollada en este trabajo.

Luego se mostraba la estructura del software, describiendo el procesado de los datos, las partes del programa, como se ejecutaban y las salidas que se reciben,

Por último se mostraba como se implementaba en la aplicación de GRADHOC.

## Técnicas estadísticas usadas

En este capítulo se comentaban las técnicas estadísticas utilizadas. Se mencionaban primero las transformaciones de los datos. Luego se hablaba de las métricas usadas para comprobar el buen funcionamiento del software. A continuación se describían las técnicas de Machine Learning que aparecen en el programa. Finalmente, se explican con detalle los tipos de gráficos de control usados y su construcción.

## Herramientas usadas y librerías de Python

En esta parte del trabajo se hablaba de todas las tecnologías, herramientas, lenguajes de programación y programas que permitían funcionar al software.

Se hablaba de las bases de datos usadas, los entornos de programación y los repositorios usados.

Finalmente se hacía una mención y una explicación de las principales librerías de Python utilizadas.

## Aplicación a datos reales

En este capítulo se hacía una revisión de los resultados obtenidos en los 3 meses de prueba del programa en diferentes instalaciones frigoríficas.

Se hablaba de los tipos de detecciones captadas, su significado y relevancia. Especificando las funcionalidades del software en tema de detección y su significado en máquinas reales.

En lo referente a las detecciones y los valores reales observados, se comentaban la relación que había entre las mayores anomalías detectadas y fallos en maquinaria producidos.



# Conclusiones y Líneas Futuras

Por último se añade partes del último capítulo del trabajo de fin de máster.

## Conclusión

En los capítulos anteriores del TFM completo se muestra el desarrollo, las principales funciones, la teoría estadística, las herramientas usadas y los resultados más relevantes de la presente propuesta para el control de máquinas, mantenimiento predictivo y detección de anomalías en la industria de la refrigeración, que incluye el diseño de la metodología y su implementación en la empresa a través de software propio.

Se ha propuesto una metodología para el control online de instalaciones de refrigeración a partir de grandes bases de datos multivariantes proporcionadas en tiempo real por plataformas web. Este procedimiento combina el ajuste de modelos de regresión propios del ML como son los árboles de decisión y el boosting, cuyas predicciones (una vez definidas las variables-características respuesta,  $Y$ , y predictores,  $X$ ) definen el funcionamiento normal de las instalaciones, con la aplicación de gráficos de control con memoria, que permiten identificar todo tipo de desviaciones con respecto al funcionamiento correcto del sistema.

Además, en este trabajo se ha interactuado intensamente con los diversos departamentos de la empresa COFRICO, conforme se ha podido realizar una clasificación de anomalías según tipo y nivel de gravedad de gran utilidad para la compañía, pues permite realizar un mantenimiento predictivo de las máquinas de refrigeración en funcionamiento en diferentes edificaciones (centros comerciales, oficinas, etc.) situadas en muy diversas localizaciones geográficas, tanto en España como en otros países.

Es importante destacar que el presente trabajo es un TFM de modalidad B, enfocado a la resolución de problemas en la empresa. En este marco, aunque en una primera instancia el trabajo a realizar en COFRICO y GRADHOC no estaba del todo definida, gracias a los conocimientos aprendidos en el Máster y por la apuesta por el desarrollo I+D+i de la empresa, se ha conseguido desarrollar un programa funcional para la detección de fallos en todo tipo de maquinaria, cuya implantación en la empresa aportará valor.

Aunque quedan etapas de desarrollo, el funcionamiento del programa es prometedor y la empresa ha empezado en este momento su implementación en todas sus instalaciones.

## Líneas futuras

Por último, en esta sección se hablaba de los desarrollos actuales del programa y los trabajos a futuro que no ha dado tiempo a implementar en los meses de desarrollo. Se mencionan solo algunas de ellas.

Primero se hará la implantación del software en todas las máquinas que intervienen en una instalación, dado que actualmente solo se está usando para controlar evaporadores. Luego el siguiente paso es adaptarlo para compresores, condensadores y válvulas de expansión.

Otra futura línea de trabajo será explorar la implantación de técnicas de predicción para series temporales, con la intención de predecir las grandes desviaciones para anticiparse a los fallos.

La última línea ya comentada por algún responsable de la empresa, será personalizar el código para que los clientes escojan que variable o variables les interesa más controlar y ellos mismos hagan la exploración, el entrenamiento y el calibrado.

Finalmente, tanto los tutores del Máster en Técnicas Estadísticas como la empresa, expresaron su interés en que se realicen parte de estas acciones e igual otras a mayores; mediante la realización de una tesis doctoral, que implicaría la realización de un doctorado en estadística industrial y una nueva relación entre la Universidade da Coruña y Cofrico.