

Detección de eventos anómalos en un entorno industrial mediante el uso de técnicas de Federated Learning.

**Darío Martín García Carretero**

Máster Universitario en Ciencia de Datos

Área 2

**Raúl Parada Medina**

**Jordi Casas Roma**

08/01/2020

Creative Commons License

This work is licensed under a [Creative Commons Attribution 4.0 International License](http://creativecommons.org/licenses/by/4.0/).

**FICHA DEL TRABAJO FINAL**

|  |  |
| --- | --- |
| **Título del trabajo:** | *Detección de eventos anómalos en un entorno industrial mediante el uso técnicas de Federated Learning.* |
| **Nombre del autor:** | *Darío Martín García Carretero* |
| **Nombre del consultor/a:** | *Raúl Parada Medina* |
| **Nombre del PRA:** | *Jordi Casas Roma* |
| **Fecha de entrega (mm/aaaa):** | 01/2020 |
| **Titulación:** | *Máster en Ciencia de Datos* |
| **Área del Trabajo Final:** | *Área 2* |
| **Idioma del trabajo:** | *Español* |
| **Palabras clave** | *eventos anómalos, entorno industrial, federated learning* |
| **Resumen del Trabajo (máximo 250 palabras):** *Con la finalidad, contexto de aplicación, metodología, resultados y conclusiones del trabajo.* | |
| Un evento anómalo es aquel que se produce de forma repentina y sin previsión. En un entorno industrial estos eventos (generalmente debidos a algún error) pueden provocar grandes daños por lo que su detección puede ayudar a prevenir situaciones críticas.  El objetivo del proyecto es detectar las posibles anomalías que se pudieran producir en los equipos industriales mediante el uso de un modelo de aprendizaje automático. Para enseñar a distinguir a un modelo entre un comportamiento normal y otro anormal se necesitan datos, contra más datos, mejor.  Hoy en día la mayoría de los componentes dentro de un entorno industrial están constantemente monitorizados mediante el uso de dispositivos de medición especializados. Debido a esto, dispondremos de los datos suministrados por nuestros dispositivos de medición para entrenar al modelo, pero ¿y si pudiéramos disponer de más datos? Muchos equipos industriales son de uso genérico y pueden ser usados para muchas tareas en muchas compañías diferentes. Si pudiéramos tener acceso a los datos de todos esos dispositivos podríamos crear un modelo mucho mejor. Por diversas razones a las compañías no les gusta compartir sus datos, es por este motivo por lo que en este trabajo se propone el uso del Federated Learning (FL). Con el FL se pueden construir modelos aprovechando toda la información disponible manteniendo la privacidad de los datos. | |
| **Abstract (in English, 250 words or less):** | |
| An anomalous event is one that occurs suddenly and without foresight. In an industrial environment, these events (normally due to some error) can cause great damage, so their detection can help to prevent critical situations.  The objective of the project is to detect possible anomalies that will occur in industrial equipment using a machine learning model. To teach model how to distinguish between a normal and an abnormal behaviour, data is needed, the more data, the better.  Today, most components within an industrial environment are constantly monitored using specialized measuring devices. Because of this, we will have the data provided by our measuring devices for the model, but what if we could have more data? Many industrial equipment is of generic use and can be used for many tasks in many different companies. If we could have access to the data of all those devices, we could create a much better model. For various reasons, companies do not like to share their data, which is why this work proposes the use of Federated Learning (FL). With FL you can build models taking advantage of all the available information and you can keep the privacy of the data. | |

**Índice**

[1. Introducción 1](#_Toc27551368)

[1.1 Contexto y justificación del Trabajo 1](#_Toc27551369)

[1.2 Objetivos del Trabajo 1](#_Toc27551370)

[1.3 Enfoque y método seguido 2](#_Toc27551371)

[1.4 Planificación del Trabajo 3](#_Toc27551372)

[1.5 Breve sumario de productos obtenidos 4](#_Toc27551373)

[1.6 Breve descripción de los otros capítulos de la memoria 4](#_Toc27551374)

[2. Estudio del estado del arte 5](#_Toc27551375)

[2.1 Detección de eventos anómalos 5](#_Toc27551376)

[2.2 Federated Learning 10](#_Toc27551377)

[2.3 Novedades propuestas 11](#_Toc27551378)

[3. Diseño e implementación del trabajo 12](#_Toc27551379)

[3.1 Consideraciones previas 12](#_Toc27551380)

[3.2 Preparación del entorno de trabajo 12](#_Toc27551381)

[3.3 Simulación de un entorno industrial 13](#_Toc27551382)

[3.3.1 Introducción 13](#_Toc27551383)

[3.3.2 Caracterización de la degradación de las máquinas 13](#_Toc27551384)

[3.3.3 Evolución de las condiciones operacionales 14](#_Toc27551385)

[3.3.4 Proceso de simulación 17](#_Toc27551386)

[3.3.5 Configuración de la simulación 18](#_Toc27551387)

[3.3.6 Resultados de la simulación 19](#_Toc27551388)

[3.4 Análisis de los datos 20](#_Toc27551389)

[4. Conclusiones 22](#_Toc27551390)

[4. Glosario 23](#_Toc27551391)

[5. Bibliografía 24](#_Toc27551392)

[6. Anexos 26](#_Toc27551393)

**Lista de figuras**

[Fig. 1 Los fallos no controlados pueden tener resultados catastróficos 1](file:///D:\projects\uoc\masterdatacience\doc\TFM_PEC3_v3.docx#_Toc27476153)

[Fig. 2 Planificación del trabajo 4](#_Toc27476154)

[Fig. 3 IRESE framework 5](#_Toc27476155)

[Fig. 4 Arquitectura de la red propuesta 6](#_Toc27476156)

[Fig. 5 Proyecciones PCA de los datos temporales 7](#_Toc27476157)

[Fig. 6 Arquitectura GAN propuesta 7](#_Toc27476158)

[Fig. 7 Modelo híbrido propuesto 8](#_Toc27476159)

[Fig. 8 Identificación de una instancia normal (izquierda) y una anormal (derecha) 9](#_Toc27476160)

[Fig. 9 Arquitectura de una Replicator Neural Network (izquierda) y función de activación de la capa intermedia (derecha) 9](#_Toc27476161)

[Fig. 10 Cada dispositivo personaliza el modelo localmente (A). Las actualizaciones de muchos usuarios se agregan (B) para actualizar el modelo global que se compartirá con todos los usuarios (C), después de lo cual se repite el proceso. 10](#_Toc27476162)

[Fig. 11 Arquitectura de la U-Net utilizada 11](#_Toc27476163)

[Fig. 12 Curvas de degradación para diferentes parámetros 14](#_Toc27476164)

[Fig. 13 Efecto de la degradación en la velocidad 15](#_Toc27476165)

[Fig. 14 Efecto de la degradación en la temperatura 16](#_Toc27476166)

**Lista de ecuaciones**

(Ecuación 1) 13

(Ecuación 2) 15

(Ecuación 3) 15

*(Ecuación 4)* 16

(Ecuación 5) 16

**Lista de tablas**

[Tabla 1 Fragmento de archivo de telemetría 19](#_Toc27497229)

[Tabla 2 Fragmento del registro de mantenimiento 19](#_Toc27497230)

[Tabla 3ejemplo 20](#_Toc27497231)

# 1. Introducción

## 1.1 Contexto y justificación del Trabajo

Un evento anómalo es aquel que se produce de forma repentina y sin previsión. En un entorno industrial estos eventos (generalmente debidos a algún error) pueden provocar grandes daños por lo que su detección puede ayudar a prevenir situaciones críticas.

El objetivo del proyecto es detectar las posibles anomalías que se pudieran producir en los equipos industriales mediante el uso de un modelo de aprendizaje automático. Para enseñar a distinguir a un modelo entre un comportamiento normal y otro anormal se necesitan datos, contra más datos, mejor. Hoy en día la mayoría de los componentes dentro de un entorno industrial están constantemente monitorizados mediante el uso de dispositivos de medición especializados. Debido a esto, dispondremos de los datos suministrados por nuestros dispositivos de medición para entrenar al modelo, pero ¿y si pudiéramos disponer de más datos? Muchos equipos industriales son de uso genérico y pueden ser usados para muchas tareas en muchas compañías diferentes. Si pudiéramos tener acceso a los datos de todos esos dispositivos podríamos crear un modelo mucho mejor. Por diversas razones a las compañías no les gusta compartir sus datos, es por este motivo por lo que en este trabajo se propone el uso del Federated Learning (FL). Con el FL se pueden construir modelos aprovechando toda la información disponible manteniendo la privacidad de los datos.

Fig. 1 Los fallos no controlados pueden tener resultados catastróficos

## 1.2 Objetivos del Trabajo

Debido a la dificultad para obtener un conjunto de datos real por ser este tipo de datos muy sensible para las compañías no se utilizarán datos reales. En su lugar, se utilizará un software que nos permitirá la simulación de las instalaciones industriales que necesitemos.

Nuestro objetivo es la creación de un modelo de Machine Learning compatible con el uso del Federated Learning (En principio cualquier algoritmo basado en algún tipo de actualización de sus parámetros es compatible con el uso del FL [1]) que sea capaz de detectar de forma eficaz los eventos anómalos producidos en los entornos industriales simulados.

## 1.3 Enfoque y método seguido

El método que se seguirá se describe a continuación:

**Preparación del entorno de trabajo**

Se instalarán las herramientas necesarias para la implementación de la simulación, la gestión del conjunto de datos y para la creación del modelo.

**Simulación de un entorno industria**

Se implementará un software que permitirá simular un entorno industrial. De esta simulación será de donde se obtendrán los datos necesarios para la creación del modelo.

**Análisis de los datos generados**

**Elección del framework Federated Learning**

Actualmente existen dos alternativas: TensorFlow y PySyft (basada en Torch) se evaluará cual es la más conveniente y se instalará el framework elegido en el entorno de trabajo.

**Elección del modelo a utilizar**

Posteriormente se decidirá sobre el tipo de modelos que se usarán basándonos en la información obtenida del análisis del estado del arte.

**Selección de los criterios de evaluación del modelo**

Se seleccionarán los métodos más convenientes para el caso de estudio (Precision, F-Score, Accuracy, etc.).

**Creación y evaluación de los modelos**

Se crearán tres tipos de modelos en función del conjunto de datos del que se disponga:

* Un modelo con todos los datos disponibles, simulando la existencia de un repositorio central para todos los datos.
* Varios modelos utilizando únicamente un pequeño conjunto de datos, simulando la existencia de “islas de información”.
* Un modelo Federated Learning con todos los datos de forma distribuida.

Una vez construidos estos modelos se analizará y se compararán sus rendimientos.

**Discusión de los resultados y análisis de posibles extensiones**

## 1.4 Planificación del Trabajo

La planificación temporal del trabajo es la siguiente:

* **18 septiembre – 20 octubre**:
  + Definir el alcance del TFM.
  + Decidir el conjunto de datos a usar.
  + Decidir el tipo de modelos/ técnicas a usar en base al estudio del estado de arte.
* **21 octubre – 21 diciembre:**
  + Instalación del entorno de trabajo
  + Obtención del conjunto de datos
  + Procesado del conjunto de datos
  + Generación de los modelos
  + Análisis de la calidad de los modelos
  + Análisis de los resultados y estudio de posibles extensiones
* **22 diciembre – 8 enero:**
  + Redacción de la memoria
  + Creación de la presentación
* **9 enero – 22 enero:**
  + Defensa pública del trabajo

A continuación, se muestra un diagrama con la programación establecida:



Fig. 2 Planificación del trabajo

## 1.5 Breve sumario de productos obtenidos

No hay que entrar en detalle: la descripción detallada se hará en el resto de capítulos.

## 1.6 Breve descripción de los otros capítulos de la memoria

Explicación de los contenidos de cada capítulo y su relación con el trabajo en global.

# 2. Estudio del estado del arte

Este trabajo aborda dos temáticas que, aunque relacionas entre sí, puede ser estudiadas de forma independiente. Por este motivo se realizará el estudio del arte de manera independiente.

## 2.1 Detección de eventos anómalos

Un evento anómalo es una observación (o pequeño conjunto de observaciones) que ocurre infrecuentemente y que se desvía o es inconsistente con respecto al resto de observaciones lo suficiente como para indicar que se está produciendo una anomalía en el conjunto de datos.

La detección de este tipo de eventos se ha estudiado ampliamente en la literatura científica desde multitud de enfoques diferentes. A continuación, se muestran algunos de ellos:

**IRESE: An intelligent rare-event detection system using unsupervised learning on the IoT edge** [2]

Este documento propone un sistema de detección de eventos anómalos basado en el uso de modelos de Machine Learning no supervisados (clustering) en tiempo real. El procedimiento se divide en cuatro fases:

1. Captura de los datos.
2. Extracción de características.
3. Clustering online, se aplica el algoritmo BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies) para generar micro-clusters.
4. Clustering offline, se utiliza Agglomerative Clustering que finalmente genera dos macro-clusters, uno que contiene los eventos normales y otro con los eventos raros.



Fig. 3 IRESE framework

Las principales ventajas que ofrece este framework son dos, una es que gran parte del proceso se lleva a cabo en dispositivos locales (edge devices) lo que reduce la sobrecarga en la red y la latencia del sistema. La otra es que no se necesita etiquetar los eventos anómalos debido al uso de algoritmos no supervisados. En cuanto a las desventajas, la principal sería la extracción de características ya que este procedimiento es muy dependiente del tipo de datos que se quieren procesar.

En cuanto a los resultados, el articulo menciona que se han obtenido valores de precision y de recall por encima del 90% en algunos casos.

**Anomaly detection in ECG time signals via deep long short-term memory networks** [3]

Este trabajo utiliza la detección de eventos anómalos con fines médicos. En particular se analiza el uso de redes neuronales LSTM (Long Short-Term Memory) profundas para la detección de anomalías en las señales EGC (electrocardiogramas).



Fig. 4 Arquitectura de la red propuesta

Si denotamos una señal EGC como X = {x(0), …, x(i), …, x(n)} donde cada x(i) representa un valor del EGC en el tiempo i. El procedimiento a grandes rasgos es el siguiente:

1. Se entrena la red con el siguiente objetivo: predecir los k siguientes valores del EGC para cada x(i).
2. A partir de la red construida se obtiene un conjunto de vectores de error para cada x(i). Con estos vectores de error se ajusta una distribución normal multivariante.
3. A partir de esa distribución se calcula la probabilidad de ocurrencia del vector de error obtenido y en función de un threshold se determina si ese punto anómalo o no.

Este método presenta varias ventajas sobre otros métodos:

1. No se requiere apenas preprocesamiento de los datos.
2. No se necesita saber a priori el tipo de anomalías que pueden aparecer.

El método proporciona resultados bastante buenos llegando valores para el

F-score por encima de 0.9 para varias clases de anomalías.

**Diagnosing Network-Wide Traffic Anomalies** [4]

El documento explora el uso del método PCA (Principal Analysis Components) para la detección de eventos anómalos en una red de comunicación. La idea consiste en dividir el espacio original en dos subespacios, uno que captura el comportamiento normal y otro el anómalo. El espacio normal estará formado por las componentes que capturan más varianza mientras que el espacio anómalo estará formado por los componentes que capturan menos varianza.



Fig. 5 Proyecciones PCA de los datos temporales

Gracias a esta división podremos descomponer cada instancia en sus componentes normales y anormales. Para determinar cuándo se ha producido un evento extraño se utiliza la norma del vector del espacio anómalo. Si esta norma no sobrepasa cierto threshold se considerará como un evento normal en otro caso se identificará como un evento anómalo.

Los resultados muestran altas tasas de detección y bajas tasas de falsos positivos.

**Anomaly Detection with Generative Adversarial Networks for Multivariate Time Series** [5]

Este estudio utiliza una arquitectura GAN (Generative Adversarial Network) como mecanismo para la detección de eventos anómalos en un entorno industrial informatizado. Una arquitectura GAN está pensada como un sistema generativo (dada una distribución, el sistema es entrenado para generar datos que sigan esa distribución). Un sistema GAN esta formado por dos redes neuronales: una llamada generativa que es la encargada de generar datos que sigan la distribución dada y una llamada discriminante, que se encarga de ayudar al entrenamiento de la primera. El objetivo del método propuesto es entrenar la arquitectura GAN para que aprenda a generar situaciones de uso normal. A continuación, se muestra un esquema del sistema propuesto:



Fig. 6 Arquitectura GAN propuesta

Para determinar cuando una situación es anómala el sistema se basa en una puntuación (una combinación entre en la diferencia entre el input real y el generado y el valor de salida de la red discriminante).

Una de las principales ventajas de este método es que es no supervisado por lo que no es necesario el etiquetado de las muestras. Por otro lado, su principal desventaja es su alto coste computacional. Con respecto a su rendimiento, el documento menciona que el modelo mejora el desempeño de modelos anteriores en el conjunto de datos utilizado (Secure Water Treatment Testbed).

**High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning** [6]

Los algoritmos one-class SVM (1SVM) son una herramienta que demostrado ser efectiva en la tarea de la detección de valores anómalos. El problema con este tipo de modelos es que no escalan bien. En este trabajo se propone un método que permite solventar esa limitación gracias al uso de DBNs (Deep Belief Networks). La idea es poder sustituir los kernels no lineales de las 1SVM (muy costosos de calcular) por kernels lineales (menos costosos) gracias al uso de las DBN para la generación de vectores de características.



Fig. 7 Modelo híbrido propuesto

Los resultados experimentales demuestran que el modelo híbrido mejora a los modelos 1SVMs en aproximadamente un 20% en algunos conjuntos de datos. Comparado con los métodos basados en autoencoders la diferencia es prácticamente inapreciable sin embargo el método propuesto es tres veces más rápido en el entrenamiento y hasta 1000 veces más rápido en el testing.

**Isolation-Based Anomaly Detection** [7]

Lo que diferencia a esta propuesta del resto es que, en lugar de tratar de modelar las instancias normales para luego buscar los valores anómalos, el modelo explícitamente identifica las anomalías. Esto se consigue a través de los llamados Isolation Forests. Los Isolation Forest están formados por un conjunto de árboles de decisión. Cada árbol se crea de la siguiente manera, primero se selecciona un atributo aleatoriamente y luego se establece una partición también aleatoria entre el máximo y el mínimo del atributo seleccionado. La idea detrás de este algoritmo es que los puntos que están más aislados estarán en las hojas mas cercanas a nodo raíz.



Fig. 8 Identificación de una instancia normal (izquierda) y una anormal (derecha)

Este algoritmo puede tratar con conjuntos de datos grandes y de alta dimensionalidad. Sin embargo, este método presenta ciertos problemas debido a que las particiones se realizan únicamente de manera vertical u horizontal. Para solventar esos problemas se ha propuesto una modificación al algoritmo original en [8] llamado Extended Isolation Forest.

**Outlier Detection Using Replicator Neural Networks** [9]

La Replicator Neural Network utilizada en este artículo es MLP (Multi-Layer Perceptron) compuesto por tres capas ocultas. La función de esta red es la de reproducir los datos de entrada en la salida con el mínimo error posible.



Fig. 9 Arquitectura de una Replicator Neural Network (izquierda) y función de activación de la capa intermedia (derecha)

Una vez entrenada la red a cada punto se le asigna un valor que es la media de los errores de reconstrucción de cada una de las variables. Cuanto más grande sea el error más probable será que se trate de una anomalía. Este método es similar al uso de auto-encoders, la diferencia más llamativa es que este método utiliza una función de activación especial en forma de escalera la capa intermedia.

## 2.2 Federated Learning

El campo del Federated Learning es relativamente nuevo y se encuentra en pleno estudio es por este motivo que no existe tantos trabajos como en el caso del Machine Learning tradicional. Se han seleccionado tres trabajos, dos de definición e introducción a esta tecnología y uno centrado en un caso práctico de aplicación.

**Communication-Efficient Learning of Deep Networks from Decentralized Data** [10]

En este artículo se presenta un algoritmo que permite el uso de modelos de Machine Learning sobre conjuntos de datos descentralizados. En el documento no solo demuestra que esta idea es factible técnicamente presentando un algoritmo llamado *FederatedAveraging* sino que permite crear modelos de una alta calidad. Uno de los aspectos más importante de esta propuesta es que permite incrementar el nivel de privacidad de los clientes ya que en ningún momento se comparten datos entre clientes ni estos son enviados a un servidor central. A continuación, se muestra un esquema del modelo propuesto:



Fig. 10 Cada dispositivo personaliza el modelo localmente (A). Las actualizaciones de muchos usuarios se agregan (B) para actualizar el modelo global que se compartirá con todos los usuarios (C), después de lo cual se repite el proceso.

**Federated Machine Learning: Concept and Applications** [11]

Este documento presenta los conceptos básicos del Federated Learning sus principales técnicas y arquitecturas. Expone su potencial en varias aplicaciones y compara sus características con otros sistemas de aprendizaje distribuidos (Distributed Machine Learning, Federated Database Systems, etc.). Adicionalmente introduce el concepto de Data Alliance of Enterprises. Una Data Alliance of Enterprises es un conjunto de organizaciones que, mediante el uso de sus datos (sin compartirlos directamente con los demás miembros) colaboran para crear un modelo que aporte beneficios a todos los participantes. Este tipo de alianzas están llamadas a ser el futuro del Machine Learning y se espera constituyan todo un nuevo modelo de negocio [12].

**Multi-Institutional Deep Learning Modeling Without Sharing Patient Data: A Feasibility Study on Brain Tumor Segmentation** [13]

En este documento se presenta el primer uso del Federated Learning para el tratamiento de imágenes para el diagnóstico médico. En particular se usa esta tecnología para la creación de un modelo distribuido de segmentación de imágenes (U-Net) sin compartir datos de pacientes entre instituciones.



Fig. 11 Arquitectura de la U-Net utilizada

El estudio demuestra que el modelo basado en Federated Learning consigue un rendimiento prácticamente equivalente al modelo creado con el uso de datos compartidos.

## 2.3 Novedades propuestas

Aunque la detección de eventos anómalos ha sido estudiada en profundidad desde múltiples perspectivas nunca se había abordado el problema de forma colaborativa y descentralizada. Esta es la principal novedad que pretende aportar este trabajo.

# 3. Federated Learning

# ¿?. Descripción del escenario

# 4. Diseño e implementación del trabajo

## 4.1 Consideraciones previas

El objetivo del trabajo es mostrar la posible aplicación del aprendizaje automático federado en la detección de potenciales fallos en máquinas dentro de un entorno industrial monitorizado. Este tipo de entornos monitorizados hoy en día están muy extendidos gracias al auge de la Industria 4.0 y la implementación de dispositivos IoT (Internet of Things) en las fábricas. Sin embargo, como ya se mencionó en apartados anteriores es prácticamente imposible obtener un conjunto de datos real por ser este tipo de datos muy sensible para las compañías. Por este motivo no se utilizarán datos reales y en su lugar, se ha desarrollado un software que nos permitirá la simulación de estos datos.

## 4.2 Preparación del entorno de trabajo

La mayor parte del desarrollo del proyecto se llevará a cabo con el lenguaje de programación Python. Python ofrece, entre otras, las siguientes ventajas [14]:

* **Simplicidad.** Python ofrece la posibilidad de desarrollar programas muy potentes con muy pocas de líneas de código. Identifica y asocia automáticamente los tipos de datos y, en general, resulta un lenguaje fácil de usar y no se requiere mucho tiempo de codificación.
* **Compatibilidad.** Muchas de las tecnologías actuales relacionadas con el Machine Learning están pesadas para ser utilizadas con este lenguaje.
* **Facilidad de aprendizaje.** En comparación con otros lenguajes, Python es fácil de aprender incluso para los programadores con menos experiencia. Principalmente por tres razones:
  + Cuenta con amplios recursos de aprendizaje.
  + Garantiza un código legible.
  + Se rodea de una gran comunidad.
* **Variedad de paquetes.** Python tiene un poderoso conjunto de paquetes para una amplia gama de necesidades de análisis y ciencia de datos ( numPy, Pandas, Scipy, Scikit-learn, etc.)
* **Visualización de datos.** Aunque no es el mejor lenguaje en lo que respecta a la visualización de datos. Actualmente existen APIs que pueden ofrecer resultados bastante buenos.

Nuestro entorno de trabajo estará formado por las siguientes herramientas:

* **Microsoft Excel.** Software de hojas de cálculo
* **Python.** En su versión 3.6
* **PyCharm.** Editor para el lenguaje de programación Python

Poner las librerías que se van a usar

## 4.3 Simulación de un entorno industrial

### 4.3.1 Introducción

El objetivo de la implementación de este software es la simulación de un entorno industrial cuyas instalaciones se encuentran monitorizadas mediante dispositivos IoT. Debido a las características únicas de cada tipo de instalación no nos es posible construir un software que simule cada una de las diferentes máquinas que pudieran existir en dicho entorno industrial. Por lo tanto, consideraremos únicamente un tipo de máquina: máquinas rotatorias genéricas. Basándonos en conjuntos de datos ya existentes (por ejemplo, Turbofan Engine Degradation Simulation Data Set) para cada máquina se ha decidido estudiar las siguientes variables operacionales[[1]](#footnote-1):

* Velocidad rotacional
* Temperatura
* Presión

Además de estos valores se tendrán en cuenta variables relacionadas con el entorno. En particular la temperatura del entorno y la presión atmosférica.

### 4.3.2 Caracterización de la degradación de las máquinas

Con el propósito de una generación de datos inspirada en la física se utilizará una aproximación similar a la utilizada en [15]. En nuestro caso se usará la siguiente ecuación para simular la degradación de una máquina en función de su tiempo de uso:

(Ecuación 1)

Donde, *a* y *b* son coeficientes que determinan la forma de la curva y *d* es la degradación inicial. A continuación se muestra un ejemplo de estas curvas:



Fig. 12 Curvas de degradación para diferentes parámetros

Podemos observar que la curva punteada representa una degradación repentina mientras que la curva continua representaría una degradación más constante ( aunque el fallo se produciría antes).

En nuestro caso se considerarán dos curvas de degradación para cada máquina lo que equivaldría a decir que cada máquina puede fallar por dos motivos diferentes:

* Uno estrechamente relacionado con la presión y/o la temperatura (a esta curva la llamaremos ***h1).***
* Una relacionada con la velocidad (que denotaremos por ***h2****)*

### 4.3.3 Evolución de las condiciones operacionales

Para cada uno de los parámetros operacionales se propondrá una fórmula[[2]](#footnote-2) que nos proporcionará su valor en el tiempo *t* en función de:

* Su valor en *t-1*
* Valores del resto de parámetros operacionales
* Valor de la(s) curva(s) de degradación en *t*
* Parámetros ambientales

Formalmente:

(Ecuación 2)

Donde:

* **F(t)** es el valor objetivo
* **w** valores de los parámetros operacionales actuales
* **F(t-1)** es el valor anterior
* **h\*(t)** valor de la(s) curva(s) de degradación
* **a\*(t)** condiciones ambientales

**Cálculo de la velocidad**

La velocidad en *t* será una media entre a velocidad en *t-1* y la velocidad objetivo ponderada con un coeficiente relacionado con la función de degradación***h2.***

(Ecuación 3)

A continuación, se muestra el efecto de la degradación en la evolución de la velocidad:

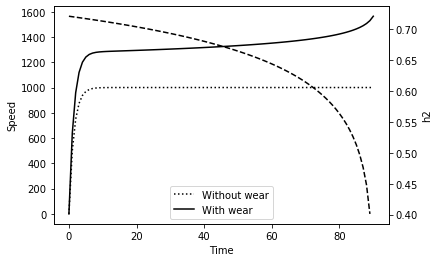


Fig. 13 Efecto de la degradación en la velocidad

En este caso, en el que la velocidad objetivo era 100, podemos ver como el efecto del desgaste afecta a la velocidad real. Al principio la diferencia se mantiene más o menos constante empezando a aumentar de forma notable al final de la vida útil de la máquina.

**Cálculo de la temperatura**

La temperatura del dispositivo estará afectada por la velocidad por lo que en este caso la temperatura estará influenciada tanto por ***h1*** como por ***h2*** (a diferencia de lo que ocurría con la velocidad a la que solo influía ***h2*** ). La fórmula utilizada para el cálculo de la temperatura es la siguiente[[3]](#footnote-3):

*(Ecuación 4)*

Donde:

(Ecuación 5)

Seguidamente se muestra gráficamente el efecto del desgaste de los componentes en la temperatura de operación:

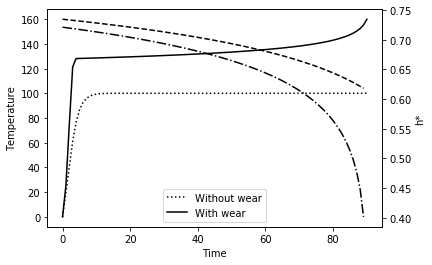


Fig. 14 Efecto de la degradación en la temperatura

Como sucedía en el caso de la velocidad vemos que la diferencia entre el caso ideal (línea punteada) y del afectado por el desgaste (línea continua) crece de forma moderada hasta que, al final de la vida útil del dispositivo, empieza a crecer de forma más pronunciada.

**Cálculo de la presión**

Igual que con la temperatura el cálculo de la presión estará influenciado por ***h1*** y ***h2.*** A continuación se muestra la fórmula usada para el cálculo de la presión[[4]](#footnote-4):

(Ecuación 6)

Donde ***g*** es la función definida en (Ecuación 5).

A continuación, se puede observar gráficamente el efecto del desgaste sobre la presión:

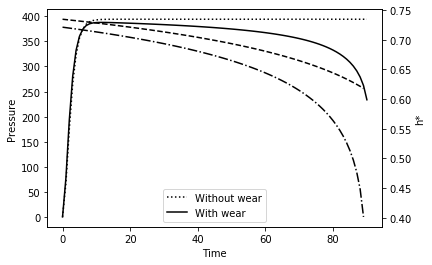


Fig. 15 Efecto de la degradación en la presión

A diferencia de lo que ocurría con la velocidad y la temperatura cuando el aparato se encuentra al final de su vida útil el valor real tiende a ser menor que el del funcionamiento óptimo.

### 4.3.4 Proceso de simulación

El objetivo es simular el funcionamiento de un conjunto de máquinas dentro de una instalación industrial y almacenar tanto los datos de telemetría como los datos referentes al mantenimiento de las máquinas.

Antes de iniciar la simulación a cada una de las máquinas que forma parte de la instalación se les asignan unas curvas de degradación (***h1***, ***h2***) cuyos parámetros son elegidos de forma aleatoria dentro de un intervalo de valores previamente establecido.

Una vez asignadas las curvas se inicia la simulación. Cada simulación (que tiene una duración determinada) se divide en lo que denominaremos periodos. En cada periodo se elige de forma aleatoria un conjunto de máquinas. Estas funcionaran durante un periodo de tiempo elegido al azar dentro de cada periodo. La velocidad de funcionamiento de cada máquina se establece aleatoriamente dentro unos márgenes establecidos. Durante su funcionamiento se almacenarán sus datos telemétricos. Puede ocurrir, que durante su funcionamiento, en una máquina se produzca un fallo. En este caso esta máquina quedará inutilizada hasta el siguiente periodo y se creará un registro de mantenimiento indicando cuando se produjo el fallo y cuál fue la causa del fallo. Una vez finalice el periodo actual esa máquina será “reparada” asignándole un nuevo par de curvas de degradación pudiendo participar nuevamente en la selección para el siguiente periodo.

El procedimiento anteriormente descrito se repite hasta que el tiempo de simulación finaliza.

### 4.3.5 Configuración de la simulación

Cada simulación tiene un total de 20 parámetros configurables. Esta capacidad de personalización nos será de gran utilidad para las siguientes fases del proyecto. Algunos de estos parámetros son:

* Duración del periodo (batch\_time)
* Duración de la simulación (simulation\_time)
* Número de máquinas en la instalación (machine\_count)
* Número de máquinas por periodo (machines\_per\_batch)
* Temperatura ambiente (temperature)
* Intervalo de velocidad de funcionamiento (operational\_speed\_min, operational\_speed\_max)

[CONFIGURATION]  
  
name = Facility A  
; three months  
simulation\_time = 7776000  
temperature = 15  
pressure = 98  
cycle\_length\_min = 1  
cycle\_length\_max = 5  
cycle\_duration = 60  
machines\_per\_batch = 5  
operational\_speed\_min = 800  
operational\_speed\_max = 900  
batch\_time = 3600  
machine\_count = 10  
ttf1\_min = 5000  
ttf1\_max = 50000  
ttf2\_min = 500  
ttf2\_max = 90000  
temperature\_max = 100  
pressure\_factor = 1.8  
telemetry\_path = ../data/generation/facility A/telemetry  
event\_path = ../data/generation/facility A/event

Fig. 16 Ejemplo de un archivo de configuración de una simulación

### 4.3.6 Resultados de la simulación

Una vez finalizada la simulación por cada máquina obtendremos dos tipos de archivos:

**Datos de la telemetría**

Como su nombre sugiere contiene los datos registrados durante el funcionamiento de la máquina. La información contenida en este archivo es la siguiente:

* **id:** identificador único de la máquina
* **time:** marca temporal de la medida
* **speed:** velocidad medida
* **tgtspeed:** velocidad objetivo
* **temp:** temperatura de la máquina
* **press:** presión en la máquina
* **atemp:** temperatura ambiente
* **apress:** presión atmosférica

Tabla 1 Fragmento de archivo de telemetría

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **id** | **time** | **speed** | **tgtspeed** | **temp** | **press** | **atemp** | **apress** |
| 2e8a9c8c | 1125 | 488.79 | 873.00 | 18.22 | 98.71 | 15.08 | 98.06 |
| 2e8a9c8c | 1126 | 727.72 | 873.00 | 22.14 | 286.18 | 15.04 | 97.94 |
| 2e8a9c8c | 1127 | 848.98 | 873.00 | 26.85 | 482.01 | 14.93 | 97.98 |
| 2e8a9c8c | 1128 | 911.73 | 873.00 | 32.51 | 608.00 | 14.94 | 97.90 |
| 2e8a9c8c | 1129 | 945.71 | 873.00 | 39.37 | 730.92 | 15.08 | 97.91 |
| 2e8a9c8c | 1130 | 953.32 | 873.00 | 47.71 | 796.18 | 14.99 | 97.96 |
| 2e8a9c8c | 1131 | 965.56 | 873.00 | 57.69 | 810.21 | 14.94 | 98.02 |
| 2e8a9c8c | 1132 | 966.21 | 873.00 | 69.78 | 834.22 | 14.95 | 98.01 |

Es importante notar que durante el proceso de simulación los datos han sido contaminados añadiendo ruido para que el proceso resultara más realista.

**Registro de mantenimiento de la máquina**

Este archivo contendrá información sobre las averías de la máquina y está compuesto por los siguientes atributos:

* **id:** identificador único de la máquina
* **time:** marca temporal del evento
* **code:** código que identifica al evento (**F1** y **F2** son los códigos de cada una de las averías que se pueden producir mientras que el código **FIXED** indica que una máquina ha sido arreglada)
* **severity: CRITICAL** si estamos ante una avería **INFO** en otro caso

Tabla 2 Fragmento del registro de mantenimiento

|  |  |  |  |
| --- | --- | --- | --- |
| **id** | **time** | **code** | **severity** |
| 2e8a9c8c | 300061 | F1 | CRITICAL |
| 2e8a9c8c | 302400 | FIXED | INFO |
| 2e8a9c8c | 2301135 | F1 | CRITICAL |
| 2e8a9c8c | 2304000 | FIXED | INFO |
| 2e8a9c8c | 2893790 | F1 | CRITICAL |
| 2e8a9c8c | 2894400 | FIXED | INFO |
| 2e8a9c8c | 4113063 | F2 | CRITICAL |
| 2e8a9c8c | 4114800 | FIXED | INFO |

## 4.4 Creación de un modelo para la detección de eventos anómalos

Antes de poder aplicar técnicas de Federated Learning es necesario construir un modelo base. Este modelo base será la piedra angular de todo el proceso de aprendizaje federado. En los siguientes apartados se detallará todo el proceso necesario para la creación de este modelo base.

### 4.4.1 Adquisición de los datos

Los datos han sido obtenidos gracias al software de simulación cuyo funcionamiento se ha detallado en el anterior. La simulación se corresponde a un periodo de temporal de aproximadamente seis meses. Los parámetros utilizados para realizar la simulación han sido:

* Número de máquinas: 20
* Número de máquinas por periodo:10
* Duración del periodo: 1h
* Temperatura ambiente media: 20 ºC
* Presión atmosférica media: 101 kPa
* Tiempo mínimo de actividad por periodo: 1 min
* Tiempo máximo de actividad por periodo: 10 min
* Mínima velocidad de funcionamiento: 800 rpm
* Máxima velocidad de funcionamiento: 900 rpm
* Temperatura máxima de funcionamiento: 125 ºC

### 4.4.2 Análisis de los datos

Debido a la forma de generar los datos de telemetría y de mantenimiento no será necesario realizar el análisis sobre todo el conjunto de datos. Se elegirá una máquina al azar y esta se considerará como representante de todo el conjunto de máquinas de la simulación. Una vez hemos seleccionado esta máquina representativa vamos a hacerle un pequeño interrogatorio a su conjunto de datos asociado:

**¿Cuál es el número de registros almacenados?**

El conjunto está formado por un total de 743144 registros

**¿Cuántos fallos se han producido en estos seis meses?**

Se han producido un total de 21 fallos de tipo **F1** y 12 de tipo **F2.** Teniendo esto en cuenta parece que son mucho más frecuentes los errores de tipo **F1**.

**¿Cuáles son sus estadísticas más relevantes?**

Tabla 3 Datos descriptivos de la máquina

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **speed** | **tgtspeed** | **temp** | **press** | **atemp** | **apress** |
| **count** | 743144 | 743144 | 743144 | 743144 | 743144 | 743144 |
| mean | 971.10 | 835.12 | 146.25 | 871.70 | 20.00 | 101.00 |
| std | 128.77 | 115.44 | 18.63 | 169.02 | 0.06 | 0.06 |
| min | 5.54 | 0.00 | 22.40 | 9.12 | 19.90 | 100.90 |
| 25% | 943.28 | 823.00 | 143.58 | 801.87 | 19.95 | 100.95 |
| 50% | 980.49 | 849.00 | 146.71 | 904.57 | 20.00 | 101.00 |
| 75% | 1014.11 | 876.00 | 151.85 | 981.73 | 20.05 | 101.05 |
| max | 1545.38 | 899.00 | 223.94 | 1753.91 | 20.10 | 101.10 |

En interesante destacar las diferencias notables que existen entre los valores medios de la velocidad (*speed*), la presión (*press*) y la temperatura (*temp*) y sus valores máximos. Si tenemos en cuenta el comportamiento descrito en la sección referida a la simulación estos valores extremos pueden estar relacionados con la aparición de fallos.

**¿Qué relación existe entre las variables operacionales?**

Tabla 4 Tabla de correlaciones

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **speed** | **tgtspeed** | **temp** | **press** |
| **speed** | 1.000 | 0.874 | 0.153 | 0.539 |
| **tgtspeed** | 0.874 | 1.000 | -0.018 | 0.387 |
| **temp** | 0.153 | -0.018 | 1.000 | -0.061 |
| **press** | 0.539 | 0.387 | -0.061 | 1.000 |

Se puede observar una correlación alta entre la velocidad (speed) y la velocidad objetivo (tgtspeed) lo cual es totalmente lógico. Sin embargo, entre el resto de las variables no se observa ninguna relación destacable.

**¿Qué aspecto tienen los datos?**

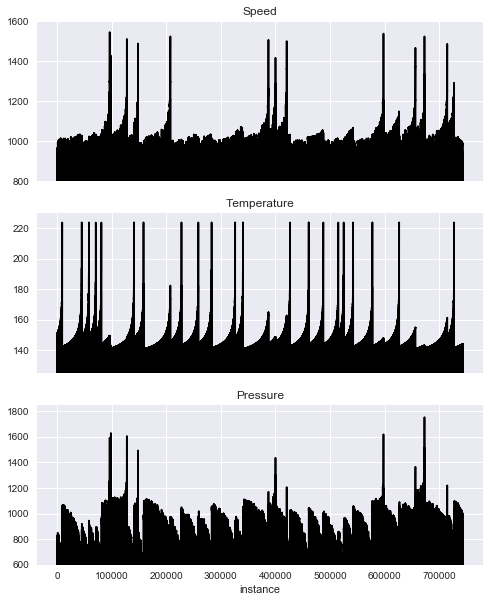


Fig. 17 Representación gráfica de los datos operacionales

Se pueden observar en todas las variables operacionales picos muy marcados. Es muy probable que estos picos estén relacionados directamente con los fallos. Para comprobar que realmente existe esa relación vamos a añadir a estas gráficas la información temporal de los fallos que se han producido:

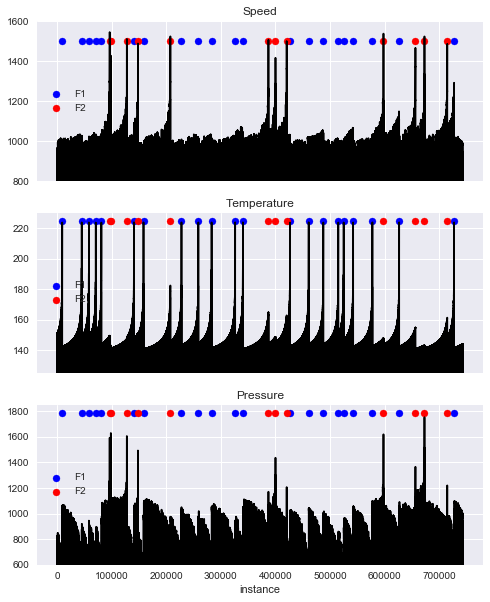


Fig. 18 Datos operacionales junto con los datos de mantenimiento

Efectivamente se puede ver que nuestra suposición era cierta en el caso de la velocidad y la temperatura. Sin embargo, en el caso de la presión se puede ver que, mientras que los fallos de tipo **F2**  sí que están relacionados con los picos los fallos de tipo **F1** están relacionados con los valles.

### 4.4.3 Elección del tipo modelo

Nuestro objetivo es poder decidir un determinado momento si una máquina está en riesgo de rotura o no, utilizando únicamente sus datos telemétricos. Dicho de otra manera, debemos clasificar esa máquina como potencialmente peligrosa o como segura. Por lo tanto parece claro que será necesario el uso de un modelo de clasificación. Existen multitud de familias de modelos que pueden ser usados para tareas de clasificación: árboles de decisión, máquinas de soporte de vectores, k-vecinos más cercanos, etc. Sin embargo nos hemos decidido por el uso de redes neuronales. Esta elección está motivada principalmente por dos cuestiones:

* Han demostrado tener un rendimiento excelente en multitud de problemas como podrían ser ……
* Aunque en teoría el uso del Federated Learning es aplicable a cualquier tipo de modelo cuyo entrenamiento se base en la optimización de parámetros , es cierto que al ser una tecnología relativamente nueva la mayoría de actuales frameworks solo permiten el uso de redes neurales en sus implementaciones de aprendizaje federado.

fallar. Eso es lo mismo que clasificar el estado de la máquina como peligroso o estado saludable.

O lo que es lo mismo clasificarla como por lo tanto el tipo de modelos

Cualquier método de optimización de parámetros

No hemos decantado por un orientar el problemas como un problema de clasificación dada una entrada en el modelo nuestro objetivo será dada una entrada clasificarla como “cercana a error F1 a la vista error F2 a la vista no errores a la vista

El planteamiento propuesto para la detección de un posible fallo es la de un método de clasificación

Existen multiples ociones para la resolución de este tipo de problemas

### 4.4.4 Procesado de los datos

Para poder entrenar y aplicar el modelo de clasificación es necesario transformarlos a un formato que los algoritmos puedan entender bla bla bla bla

Pero antes de describir el proceso que se ha llevado a cabo es necesario enterder el concepto de ciclo.

**¿Qué son los ciclos?**

Muchas máquinas del mundo real funcionan en ciclos. Un ciclo puede considerarse como un período temporal que describe un estado de funcionamiento de una máquina. Por ejemplo, la operación de un motor en un avión puede describirse por los ciclos: motor en funcionamiento (avión en vuelo) o motor apagado (avión en tierra).

**¿Porque es importante tenerlos en cuenta?**

Las transmisiones de telemetría sin procesar, si bien pueden ser muy útiles para tareas como el monitoreo en tiempo real, pueden causar problemas a la hora de construir modelos de detección de fallos ya que es frecuente que en entornos no controlados (como puede ser una fábrica) los datos no sean del todo precisos.

Para mitigar los posibles errores en las mediciones (fallos puntuales en los equipos de medida, ruido, etc.) que pueden afectar de forma dramática a la calidad de los modelos construidos suele ser una buena opción considerar datos agregados por ciclo.

.

**Elección del framework Federated Learning**

Actualmente existen dos alternativas: TensorFlow y PySyft (basada en Torch) se evaluará cual es la más conveniente y se instalará el framework elegido en el entorno de trabajo.

**Elección del modelo a utilizar**

Posteriormente se decidirá sobre el tipo de modelos que se usarán basándonos en la información obtenida del análisis del estado del arte.

**Selección de los criterios de evaluación del modelo**

Se seleccionarán los métodos más convenientes para el caso de estudio (Precision, F-Score, Accuracy, etc.).

**Creación y evaluación de los modelos**

# 5. Conclusiones

Este capítulo tiene que incluir:

* Una descripción de las conclusiones del trabajo: Qué lecciones se han aprendido del trabajo?.
* Una reflexión crítica sobre el logro de los objetivos planteados inicialmente: Hemos logrado todos los objetivos? Si la respuesta es negativa, por qué motivo?
* Un análisis crítico del seguimiento de la planificación y metodología a lo largo del producto: Se ha seguido la planificación? La metodología prevista ha sido la adecuada? Ha habido que introducir cambios para garantizar el éxito del trabajo? Por qué?
* Las líneas de trabajo futuro que no se han podido explorar en este trabajo y han quedado pendientes.

# 6. Glosario

Definición de los términos y acrónimos más relevantes utilizados dentro de la Memoria.

# 7. Bibliografía

|  |  |
| --- | --- |
| [1] | F. Hartmann. [En línea]. Available: https://florian.github.io/federated-learning/. [Último acceso: 15 10 2019]. |
| [2] | Z. Janjua, M. Vecchio, M. Antonini y F. Antonelli, «IRESE: An intelligent rare-event detection system using unsupervised learning on the IoT edge,» *Engineering Applications of Artificial Intelligence,* vol. 84, pp. 41-50, 2019. |
| [3] | S. Chauhan y L. Vig, «Anomaly detection in ECG time signals via deep long short-term memory networks,» de *IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, Paris, 2015. |
| [4] | A. Lakhina, M. Crovella y C. Diot, «Diagnosing Network-Wide Traffic Anomalies,» *Computer Communication Review,* vol. 34, 2004. |
| [5] | D. Li, D. Chen, J. Goh y S. k. Ng, «Anomaly Detection with Generative Adversarial Networks for Multivariate Time Series,» de *International Workshop on Big Data*, London, 2018. |
| [6] | S. M. Erfani, S. Rajasegarar, S. Karunasekera y C. Leckie, «High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning,» *Pattern Recognition,* vol. 58, pp. 121-134, 2016. |
| [7] | F. T. Liu, K. Ting y Z.-H. Zhou, «Isolation-Based Anomaly Detection,» *ACM Transactions on Knowledge Discovery From Data - TKDD,* vol. 6, pp. 1-39, 2012. |
| [8] | S. Hariri, M. Carrasco Kind y R. J. Brunner, «Extended Isolation Forest,» 2018. |
| [9] | S. Hawkins, H. He, G. Williams y R. Baxter, «Outlier Detection Using Replicator Neural Networks,» *Data Warehousing and Knowledge Discovery,* pp. 170-180, 2002. |
| [10] | H. B. McMahan, E. Moore, D. Ramage, S. Hampson y B. Agüera y Arcas, «Communication-Efficient Learning of Deep Networks from Decentralized Data,» 2016. |
| [11] | Q. Yang, Y. Liu, T. Chen y Y. Tong, «Federated Machine Learning: Concept and Applications,» 2019. |
| [12] | A. Gonfalonieri, «Towards Data Science,» [En línea]. Available: https://towardsdatascience.com/federated-learning-a-new-ai-business-model-ec6b4141b1bf. |
| [13] | M. J. Sheller, G. A. Reina, B. Edwards, J. Martin y S. Bakas, «Multi-Institutional Deep Learning Modeling Without Sharing Patient Data: A Feasibility Study on Brain Tumor Segmentation,» 2018. |
| [14] | U. I. d. Valencia, «www.universidadviu.es,» [En línea]. Available: https://www.universidadviu.es/python-para-big-data-motivos-para-elegirlo/. [Último acceso: 15 12 2019]. |
| [15] | K. G. D. S. N. E. A. Saxena, «Damage propagation modeling for aircraft engine run-to-failure simulation,» de *International Conference on Prognostics and Health Management,*, Denver, CO, USA, 2008. |
| [16] | Microsoft, «gallery.azure.ai,» [En línea]. Available: https://gallery.azure.ai/Collection/Predictive-Maintenance-Template-3. [Último acceso: 16 11 2019]. |

# 6. Anexos

Listado de apartados que son demasiado extensos para incluir dentro de la memoria y tienen un carácter autocontenido (por ejemplo, manuales de usuario, manuales de instalación, etc.)

Dependiente del tipo de trabajo, es posible que no haya que añadir ningún anexo.

1. Las unidades de medida serán: kPa para las presiones, ºC para las temperaturas y rpm para las velocidades [↑](#footnote-ref-1)
2. Las fórmulas propuestas están basadas en las fórmulas presentadas en [16]. [↑](#footnote-ref-2)
3. Como se verá en un apartado posterior, el valor de las constantes que aparecen en la fórmula podrá variar en función de las condiciones de la instalación que se quiera simular. [↑](#footnote-ref-3)
4. Al igual que el caso de la temperatura las constantes pueden variar en función de las condiciones de la instalación que se quiera simular. [↑](#footnote-ref-4)