



UNIVERSIDAD DE CHILE  
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS  
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA

INDICADOR PARA PRONÓSTICO DE FALLAS EN MAQUINARIA DE PLANTA DE  
ALIMENTOS

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL ELÉCTRICO

ALVARO VALENZUELA PORCILE

PROFESOR GUÍA:

SR. MARCOS ORCHARD CONCHA

MIEMBROS DE LA COMISIÓN:

SR. JORGE SILVA SÁNCHEZ

SR. ANDRES PIMENTEL ZAMORA

SANTIAGO DE CHILE

2020

RESUMEN DE LA MEMORIA PARA OPTAR  
AL TÍTULO DE INGENIERO CIVIL ELÉCTRICO  
POR: ALVARO ESTEBAN VALENZUELA PORCILE  
FECHA: 2020  
PROF. GUÍA: MARCOS ORCHARD CONCHA

## INDICADOR PARA PRONÓSTICO DE FALLAS EN MAQUINARIA DE PLANTA DE ALIMENTOS

Hoy en día la industria en la búsqueda de optimizar sus procesos y mejorar la toma de decisiones, se ha empezado a tomar mayor consciencia de los datos que poseen y buscar crear un mayor valor de ellos. En conjunto con el aumento en la capacidad de almacenaje de datos y el crecimiento en investigación sobre métodos de inteligencia computacional, las implementaciones de estas técnicas se han vuelto más comunes, siendo aplicadas en diversas áreas industriales y comerciales.

Por lo anterior es que el diagnóstico y pronóstico de fallas se ha convertido en un tema de interés para varias empresas en la búsqueda de disminuir costos económicos y logísticos en sus líneas de producción. Y gracias al aumento de investigación en algoritmos de inteligencia computacional y disponibilidad de software, la implementación de sistemas de pronóstico de fallas se ha vuelto un tema de investigación en constante crecimiento.

Este trabajo consiste en la implementación de un indicador para la degradación de la prensa peletizadora de una planta de alimento para animales, en específico para una falla de atolle de la prensa. El método a utilizar consistió en la modelación de la degradación del equipo en función del tiempo de falla y el uso de una red neuronal que entrega una estimación del estado de la maquina y con la incertidumbre asociada.

Al hacer el análisis de los resultados del estimador en diferentes situaciones utilizando el error cuadrático medio, evaluando el desempeño que tiene en función de los tiempo de falla y en función del tipo de modelo de degradación.

A través de los resultados se concluye que la propuesta de la estimación de la degradación como indicador para el pronostico de fallas es una opción viable para ciertas condiciones, pero para prepararlo para un ambiente de producción es necesario avanzar en diferentes ámbitos para obtener una mayor confiabilidad en los resultados.

*Para mi familia y amistades en el camino.*

# Tabla de contenido

<b>1. Introducción</b>	<b>1</b>
1.1. Motivación . . . . .	1
1.2. Alcances . . . . .	2
1.3. Objetivos . . . . .	2
1.3.1. Objetivos Generales . . . . .	2
1.3.2. Objetivos Específicos . . . . .	2
1.4. Estructura de la Memoria de Título . . . . .	2
<b>2. Marco Teórico</b>	<b>4</b>
2.1. Plantas de alimento para animales . . . . .	4
2.1.1. Etapa de peletizado . . . . .	5
2.1.2. Prensa de pellet . . . . .	5
2.2. Mantenimiento basado en condición . . . . .	6
2.3. Redes Neuronales . . . . .	7
2.3.1. Funciones de activación . . . . .	7
2.3.2. Redes FeedForward . . . . .	8
2.3.3. Entrenamiento . . . . .	8
2.3.4. Regularización . . . . .	8
2.3.5. Incerteza en redes neuronales . . . . .	9
2.4. Distribución Beta . . . . .	10
<b>3. Metodología</b>	<b>11</b>
3.1. Formalización del problema . . . . .	11
3.2. Descripción de los datos . . . . .	12
3.2.1. Caracterización de modo de falla a modelar . . . . .	13
3.3. Pre-procesamiento de Datos . . . . .	15
3.4. Modelo de Estimación de salud . . . . .	17
3.5. Diseño del Estimador . . . . .	18

<b>4. Análisis de Resultados</b>	<b>20</b>
4.1. Parámetros Estimador . . . . .	20
4.2. Filtrado de Característica . . . . .	20
4.3. Estimación de degradación . . . . .	21
4.3.1. Dataset $D_{>24h}$ . . . . .	21
4.3.2. Dataset $D_{<24h}$ . . . . .	24
<b>5. Conclusiones</b>	<b>25</b>
<b>Bibliografía</b>	<b>27</b>

# Índice de cuadros

4.1. Parámetros del Modelo. . . . .	20
4.2. Resultados de los Modelos. . . . .	21

# Índice de figuras

2.1. Imagen de prensa de pellet. . . . .	5
2.2. Elemento de prensa de peletizado [2]. . . . .	6
2.3. Piezas interiores de la prensa [2]. . . . .	6
2.4. Ejemplo de arquitectura de perceptron multicapa. . . . .	8
2.5. Ejemplos de distribuciones Beta con diferentes parámetros. . . . .	10
3.1. Diagrama de metodología. . . . .	12
3.2. Diagrama Sankey de las fallas registradas de las prensas de pellet. . . . .	14
3.3. Histograma de tiempos de falla. . . . .	14
3.4. Cluster de los parámetro del modelo de falla y los tiempos entre fallas correspondiente. . . . .	15
3.5. Ejemplo del uso de ventanas para creación características. . . . .	16
3.6. Función de degradación con diferentes valores de C. . . . .	18
4.1. Resultados de modelo de degradación lenta en dataset $D_{>24h}$ . . . . .	22
4.2. Comparación resultados de modelo de degradación lenta en dataset $D_{>24h}$ . . . . .	22
4.3. Comparación resultados de modelo de degradación lenta en dataset $D_{>24h}$ en tiempos límites. . . . .	23
4.4. Resultados de modelo rápida de degradación en dataset $D_{>24h}$ . . . . .	23
4.5. Resultados de modelo de degradación lenta en dataset $D_{<24h}$ . . . . .	24

# Capítulo 1

## Introducción

### 1.1. Motivación

La industria de alimentos ha sido esencial en la sociedad actual, aportando en facilitando su transporte y proporcionando una producción en masa más eficiente y económica Y no sólo para consumo humano, sino también animal, con un enfoque centrado en un mejor control del valor nutritivo de los alimentos entregados.

Debido a este crecimiento en la industria alimenticia, muchos procesos se han automatizados, convirtiendo trabajos que antes eran manuales a maquinarias dentro de una línea de producción. La implementación de los controladores y sistemas de SCADA para estas maquinarias ha provocado un aumento en la cantidad de datos disponibles que se tiene de los procesos, facilitando su control y caracterización.

Una problemática importante en cualquier proceso industrial son las fallas de maquinarias claves, lo cual comúnmente implican la detención completa o parcial de la operación, provocando un costo financiero. Comúnmente estas maquinarias pueden tener diferentes modos de fallas, con distintos efectos y causas raíces, por lo cual su estudio es una etapa importante para un correcto análisis.

Por lo anterior es que el diagnóstico y pronóstico de fallas se ha convertido en un tema de interés para varias empresas en la búsqueda de disminuir costos en sus líneas de producción. Y gracias al aumento de investigación en algoritmos de inteligencia computacional y disponibilidad de software, la implementación de sistemas de pronóstico de fallas se ha vuelto un tema de investigación en constante crecimiento.

Lo que se propone en esta Memoria de Título es la caracterización del modo de falla más frecuente de la línea de producción de alimentos para animales y la implementación de un indicador probabilístico para el pronóstico de este mismo, utilizando como datos de entrada las variables que aparecieron más significativas en su estudio.



## 1.2. Alcances

Este trabajo está basado en datos de maquinarias en plantas de alimentos para animales y fueron entregadas por la empresa WSP. Esta empresa es una consultoría multidisciplinaria de ingeniería, la cual tiene un área que está encargada de la automatización de los procesos y supervisión de control y adquisición de datos.

El propósito de este trabajo es el acercamiento a un sistema para el pronóstico de fallas para las maquinarias claves de una planta de alimentos, con el fin de poder coordinar las acciones óptimas para disminuir los efectos negativos que causan en la productividad de la empresa.

El diseño del indicador se hará en función de la caracterización del modo de falla utilizando tanto los datos entregados, el conocimiento experto de los operadores y modelos físicos de la maquinaria. Además, en la actualidad las implementaciones de estos algoritmos han tendido a modelar también la incertidumbre [1] [9] [13] [7] dando así una mayor información a la predicción posibilitando que el usuario pueda trabajar en función de intervalos de confianza. Es por eso que se desea que el algoritmo a utilizar consista en la actualización de un modelo a priori del estado de la maquinaria utilizando una red neural.

## 1.3. Objetivos

### 1.3.1. Objetivos Generales

Desarrollo de indicador probabilístico para la predicción de fallas en maquinaria crítica en planta de alimentos utilizando diferentes variables en función de los datos históricos, conocimiento experto y modelos físicos de la maquinaria.

### 1.3.2. Objetivos Específicos

- Análisis de los modos de fallas y variables relevantes del sistema.
- Integración de conocimiento experto y modelos físicos en el diseño del algoritmo.
- Definición del modelo del estado de degradación de la maquinaria.
- Implementación de red neuronal para la inferencia de indicador de falla
- Selección de características de relevancia para el indicador de falla.

## 1.4. Estructura de la Memoria de Título

La Memoria de Título consistirá de 5 capítulos, los cuales se estructuran de la siguiente manera.

En el Capítulo 1 se describe la motivación, alcances y se definen los objetivos del trabajo.

El Capítulo 2 es el marco teórico en donde se presentan los antecedentes de las plantas de alimento peletizado y también los diferentes procesos y maquinarias que funcionan dentro de ellas.

En el Capítulo 3 se explica la modelación y metodología utilizada, presentado la formalización del problema, el procesamiento de los datos y los diferentes modelos que se evaluaron.

En el Capítulo 4 se presenta los resultados y análisis del modelo seleccionado.

Finalmente, en el Capítulo 5 se entrega las conclusiones y el trabajo a futuro que se plantea en función de los resultados obtenidos.

## Capítulo 2

# Marco Teórico

### 2.1. Plantas de alimento para animales

La modernización en la producción de alimento ha sido necesaria de acuerdo al crecimiento de demanda en la sociedad actual, buscando abastecer una población en constante crecimiento, con producciones de grandes magnitudes y con procesos involucrando diferentes tipos de maquinaria.

En la búsqueda de la uniformidad y eficiencia de los procesos, en los comienzos del siglo 20, el uso de molinos martillos (para la molienda de los ingredientes) y mezcladores de batch, comenzó a ser una práctica común en los productores de alimento. Pero el mayor impacto fue la aparición del pellet en 1920 y consolidado en la industria en 1930. Este método entrega un producto cilíndrico de fácil manipulación en donde se comprime diferentes tipos de ingredientes.

Las plantas de alimentos actuales consisten de varias etapas:

- **Abastecimiento:** La recepción a silos, bodegas o tanques de los diferentes ingredientes a utilizar en los alimentos a producir.
- **Molienda:** El procesamiento de ingredientes que necesitan ser molidos para tener un tamaño y/o granulometría óptima.
- **Dosificación:** La selección de los ingredientes a utilizar en las diferentes órdenes de trabajos con sus respectivas recetas y proporciones.
- **Mezclado:** La mezcla de los diferentes ingredientes ya procesados para obtener un producto más homogéneo.
- **Peletizado:** La producción de pellet a través de la extrusión de la mezcla por un molde, la cual se le ha agregado humedad y calor, en forma de vapor.
- **Enfriado:** El enfriamiento del pellet para procesos posteriores

- **Aceitado:** En el aceitado se integra aceites con diferentes valores nutritivos al pellet. Este proceso puede ser hecho durante el peletizado, en el acondicionamiento, o después del enfriado.
- **Despacho:** Procesos necesarios para dejar el producto listo para la entrega a los clientes.

### 2.1.1. Etapa de peletizado

La etapa de peletizado es la encargada de procesar la mezcla de los ingredientes y entregar las propiedades físicas del pellet. Este proceso consiste de 2 etapas principales y una opcional.

El primer proceso es el termo-acondicionamiento, en donde la mezcla de los diferentes ingredientes pasa por el acondicionador, el cual tiene la tarea de agregar calor y agua entregando como producto una masa más uniforme.

Luego vendría el proceso mecánico de peletización, el cual consiste el paso de esta masa acondicionada por la prensa de pellet y por donde se obtiene el pellet en si. En la sección siguiente se explica en mayor detalles el funcionamiento de esta prensa.

Existe un proceso opcional intermedio, en el cual después del acondicionamiento, para ciertos alimentos, se pasa la masa por una maquinaria llamada Feed Expander (también llamado FEX), la que tiene como fin la entrega de energía a la masa en forma de compresión. Esta compresión le entrega mejores cualidades física al producto lo cual facilita su manejo en el proceso de peletización.

### 2.1.2. Prensa de pellet

Las prensas de pellet, o también conocidos como peletera, son las maquinas que reciben la mezcla después de haber pasado por el acondicionador y opcionalmente por el feed expander y presiona esta masa a través de agujeros de un molde cilíndrico, en donde se consigue la forma final del pellet. La



Figura 2.1: Imagen de prensa de pellet.

prensa consiste en 4 elementos que se pueden ver en la Figura 2.2. En la zona 1 están los dispositivos de peletizado, por el cual por la abertura superior entra la masa y por la inferior salen los pellets. La zona 2 consiste en los engranajes de reducción de velocidad, esto debido a que las velocidades de los moldes son diferentes a los motores convencionales. En la zona 3 se ubica el motor eléctrico que está montada en una base común, la cual es la zona 4.

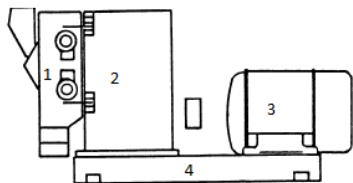


Figura 2.2: Elemento de prensa de peletizado [2].

El mecanismo de peletizado se puede ver en la Figura 2.3, en donde al entrar la masa (2.3.1) a la zona 1 de la Figura 2.2, es presionada por los rodillos fijos (2.3.4) hacia los agujeros del molde cilíndrico (2.3.5), los cuales están girando en el mismo plano vertical. En el exterior del molde se ubican cuchillas (2.3.3) las cuales cortan los pellets (2.3.2) en un tamaño adecuado.

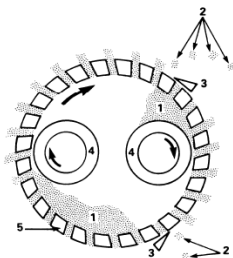


Figura 2.3: Piezas interiores de la prensa [2].

## 2.2. Mantenimiento basado en condición

El mantenimiento basado en condición (**MBC**) es el uso de los datos del estado actual de la máquina para determinar la condición de la maquinaria y así también su condición de falla, con el objetivo de programar reparaciones y acciones de mantenimiento antes de una falla.

La evaluación de la condición de la maquinaria es lograda comúnmente utilizando la información de los sensores y maquinaria de los procesos, registros de mantenimiento, métricas de calidad y contexto de operación. Luego se diseña algún indicador y en función de el se hace la toma de decisión sobre las acciones de mantenimiento correspondientes.

La estrategia de MBC es utilizadas para diferentes objetivos. Uno de los objetivos es la evaluación de condición, que comúnmente consiste en la capacidad de poder discernir de comportamientos normales y anómalos del sistema. Otro objetivo es el diagnóstico, que consiste en reconocer estados de operación y modos de fallas específicos del sistema o maquinaria en tiempo real. Y por ultimo, y en el que se esta haciendo mayor investigación, es el pronóstico, que consiste en la predicción de condición de la maquina, lo cual consiste comúnmente en la predicción de la vida útil restante (RUL, por sus siglas en ingles), la cual consiste formalmente en un problema de regresión para pronósticos

de tiempos de falla. [5]

Con estos objetivos se ha fomentado la creación de varios algoritmos, los cuales no solo se enfocan en la entrega del indicador, parámetros o predicción, sino también intervalos de confianza o la incertidumbre de los resultados, entregando mayor información para la interpretación de los resultados [11].

Los algoritmos que se utilizan en MBC se basan ya sea en modelos físicos matemáticos, utilizando el conocimiento experto del sistema y/o modelos aprendidos de los datos, lo que implica una dependencia a un gran volumen de datos. Aunque hoy en día la mayoría de ellos son una combinación del enfoque basado en modelo y en datos.

Los tipos de algoritmos que se utilizan comúnmente son a través de filtros de Kalman o partículas con modelos matemáticos físicos [14] [8] [3] o a través de métodos mas orientados a los datos como las redes neuronales (feed forward, recursivas, convolucionales, etc) [15] [12]

## 2.3. Redes Neuronales

Las redes neuronales son una modelo de inteligencia computacional basado en la conexión de varias unidades llamadas neuronas, la cuales comúnmente se agrupan en capas, en donde las salidas de las neuronas de una capa son los datos de entrada de una siguiente capa. [6]

Las neuronas se caracterizan por 2 elementos, los pesos y la función de activación.

Los pesos se definen como un vector  $\vec{W} \in \mathbb{R}^n$ , cuyos valores son ponderadores de las entradas de la neurona y son calculados a través del proceso de entrenamiento de la red, además existe el bias  $b \in \mathbb{R}$  el cual es sumado a la multiplicación de la entrada con los pesos.

La función de activación se define como una función  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  la cual define la salida de neurona con la siguiente ecuación:

$$N(\vec{X}) = \phi(\vec{W} \cdot \vec{X} + b) \quad (2.1)$$

$$= \phi\left(\sum_{i=0}^n w_i x_i + b\right) \quad (2.2)$$

Donde  $N(x)$  es la salida de la neurona,  $n$  es la dimensionalidad de  $\vec{X}$ ,  $w_i$  es el elemento i-esimo de  $\vec{W}$  y  $x_i$  es el elemento i-esimo de  $\vec{X}$ .

### 2.3.1. Funciones de activación

Las funciones de activación son utilizadas en las redes neuronales para varios propósitos, ya sea ingresar no linealidades a la función, acotar el rango de la salida o facilitar el aprendizaje de la red.

En función de las características anteriores, existe una gran variedad de funciones, a continuación, una lista de funciones de activaciones comúnmente utilizadas y con diferentes cualidades:

$$\text{Sigmoide}(x) = \frac{1}{1 + e^x} \quad (2.3)$$

$$\text{RELU}(x) = \max(x, 0) \quad (2.4)$$

$$\text{Softplus}(x) = \ln(1 + e^x) \quad (2.5)$$

### 2.3.2. Redes FeedForward

Las redes Feedforward se refieren a un estilo de arquitectura de redes neuronales que está conformada por un grafo aciclico dirigido, en otras palabras, una red neuronal que no posee ningún ciclo en donde las variables de entrada de una neurona de la red no son función de la salida de esa misma neurona.

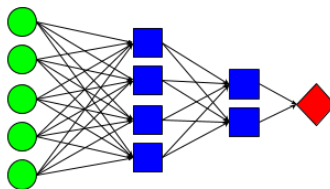


Figura 2.4: Ejemplo de arquitectura de perceptrón multicapa. Capa de entrada (Verde), capa oculta (Azul) y salida (Rojo).

Una arquitectura común de redes feedforward son los perceptrón multicapa como la que aparece en la Figura 2.4, la cual consiste en capas de neuronas secuenciales, en donde la salida de una capa es la entrada de la capa siguiente.

### 2.3.3. Entrenamiento

El entrenamiento de las redes neuronales consiste en el proceso de optimización de los valores de los pesos y bias para minimizar una función de costo  $J$ . Los valores de las redes neuronales son ajustados iterativamente a través del algoritmo de back propagation [10], en donde se calcula el gradiente del error en función de los diferentes pesos.

### 2.3.4. Regularización

Los métodos de regularización están enfocados en resolver el problema de sobre ajuste de los algoritmos agregando restricciones a las soluciones.

## Penalización de Norma

Penalización de norma consiste en agregar a la función de costo la norma de los valores de los pesos de la neurona. Los métodos más comunes consisten en utilizar la norma L1 o L2, y luego ponderarlos por un factor el cual hace el balance entre minimizar la función de costo o la regularización. En ambos casos la función de costo quedaría como a continuación:

$$J_{reg}^{L1} = J + \lambda \sum \frac{|w_i|}{N} \quad (2.6)$$

$$J_{reg}^{L2} = J + \lambda \sum \frac{\|w_i\|^2}{N} \quad (2.7)$$

Donde  $J$  es la función de costo original,  $w_i$  son en este caso los pesos de la neurona, pero también podrían ser los pesos de una regresión lineal,  $N$  es la cantidad de pesos,  $\lambda$  es un ponderador para controlar la penalización en comparación a la función de costo original y  $J_{reg}^{L1}$  y  $J_{reg}^{L2}$  son las nuevas funciones de costo.

## Dropout

El método de dropout consiste en la desactivación aleatoria de ciertas neuronas, se designa un valor  $p_{DO}$  entre 0 y 1 para cada capa de neuronas, el cual será la probabilidad de que cada neurona se desactive, lo que significa que su salida será 0. Este proceso de dropout solo se realizara en entrenamiento, para inferencia se utilizara sin esta probabilidad de desactivación.

### 2.3.5. Incerteza en redes neuronales

Es común que en la definición de problemas de regresión y clasificación se tenga un mayor enfoque en el valor de la predicción, pero hoy en día para muchas aplicaciones y ámbitos se ha vuelto una necesidad tener un intervalo de confianza o la incerteza de lo obtenido. Para lograr esto es necesario formular el problema de una forma probabilística más rigurosa, Los 2 tipos de incerteza que explicaremos son la incerteza epistémica y aleatoria [4].

La incerteza epistémica consiste en aquella que explica la confiabilidad del modelo seleccionado, lo cual es comúnmente modelada con una distribución de probabilidad de las familias de funciones evaluadas, en el caso de las redes neuronales se modela como una distribución de valores en los pesos y bias. Esta incerteza debería reducirse en función de la cantidad de datos.

La incerteza aleatoria representa el ruido de medición en el contexto, comúnmente es modelada a través de una distribución de probabilidades en la salida y en función de la verosimilitud de ella. Este tipo de incerteza no disminuye con la cantidad de datos, pero si con la calidad de datos.



## 2.4. Distribución Beta

La distribución beta es una familia de distribuciones de probabilidades continuas acotadas en un intervalo con un mínimo y máximo, comúnmente entre  $[0,1]$  o  $(0,1)$  y parametrizadas por 2 variables positivas ( $\alpha$  y  $\beta$ ). Debido a lo anterior es comúnmente utilizada para modelar variables aleatorias de problemas de diferentes disciplinas que están acotadas a cierto intervalo.

La función de densidad de probabilidades de la distribución Beta es la siguiente:

$$PDF_{Beta(\alpha,\beta)}(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}$$

donde  $B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$  y  $\Gamma$  es la función Gamma.

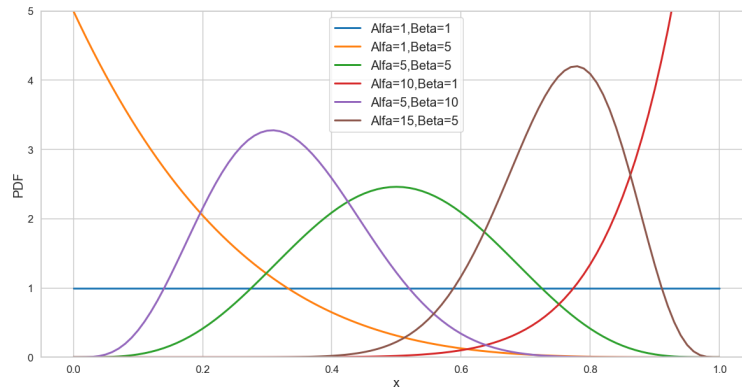


Figura 2.5: Ejemplos de distribuciones Beta con diferentes parámetros.

Las propiedades comunes de la distribución son las siguientes:

$$Media(Beta(\alpha,\beta)) = \frac{\alpha}{\alpha + \beta} \quad (2.8)$$

$$Moda(Beta(\alpha,\beta)) = \frac{\alpha - 1}{\alpha + \beta - 2} \text{ para } \alpha, \beta > 1 \quad (2.9)$$

$$Varianza(Beta(\alpha,\beta)) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)} \quad (2.10)$$

# Capítulo 3

## Metodología

### 3.1. Formalización del problema

El tema de esta Memoria de Título consistió en la implementación de un indicador de falla en una planta de alimentos para animales, de la cual se obtuvo diferentes variables de distintas maquinarias y sensores, además de registros de mantención y producción, los cuales fueron entregados a través de la empresa WSP, que está encargada de los sistemas de control de la planta de alimentos. El trabajo necesario para lograr los objetivos consiste de varias etapas.

Primero se hizo un análisis de los distintos modos de falla, analizando las causas raíces, consecuencias, frecuencia, riesgo de seguridad, componentes afectados y acciones de mantenimiento. Además, para el indicador de falla, se seleccionó el modo de falla que tuviera un efecto considerable en la producción de la planta y una frecuencia importante en comparación a las otras. Esto se logró analizando los registros de mantención y recopilando conocimiento experto de los operadores.

La segunda parte fue el pre-procesamiento de datos, en donde fue necesario analizar y formatear las diferentes fuentes de información para la creación de una base de datos unificada para un correcto entrenamiento y evaluación de los algoritmos a utilizar. Los datos de variables de máquinas, sensores e insumos de los procesos están en una base de datos de Oracle, mientras que los registros de mantención fueron entregados como un archivo Excel. Para esta etapa se utilizó el lenguaje de programación de Python para crear un dataset de series de tiempos de fallas con las variables de interés correspondientes.

Por último, es la implementación y evaluación del algoritmo para el indicador de fallas, en donde se utilizó una red neuronal cuyas variables de entradas fueron seleccionadas usando tanto el análisis de los datos históricos como el conocimiento experto recopilado.

El indicador de falla fue la regresión del modelo del estado de degradación de la maquina  $S_d$ , el cual consiste de una función creciente entre 0 y 1, donde 1 representa que la maquina está en falla y 0 después de una acción de mantención relaciona al estado de falla.

El enfoque de este indicador de falla es solo estimar el estado de salud de la maquina en función del modo de falla elegido, este indicador en una etapa posterior será utilizado para la predicción a través de algoritmo de regresión de  $S_d$  y tendrá un enfoque bayesiano, en donde se tendrá un  $S_d$  a priori el cual se actualizará con datos a posteriori utilizando los datos de la planta en una ventana de tiempo.

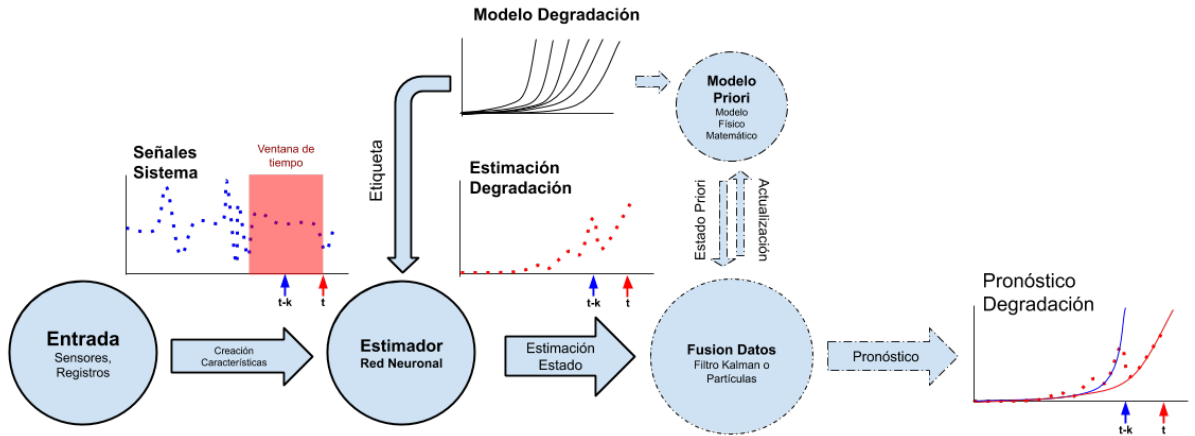


Figura 3.1: Diagrama de metodología.

El modelo a priori inicial del estado de degradación  $S_{d_{prior}}$  será solo función del tiempo desde la última falla y se parametriza utilizando los datos históricos de los tiempos entre fallas y la aproximación del modelo físico de la evolución de degradación de la máquina. Este modelo a priori dependerá de un parámetro  $\alpha$  el cual se caracterizará con una media y varianza, la cual se actualizará iterativamente utilizando datos posterioris.

Los datos posteriori en este caso son el indicador de falla, cual entregara el estado de degradación del equipo. Finalmente, con el uso de  $\alpha$  como una distribución de valores y su actualización a través de una red neuronal bayesiana que también posee incertidumbre, se obtiene un indicador de falla con estas mismas características lo cual facilita su interpretación en un ambiente de producción.

### 3.2. Descripción de los datos

Los datos utilizados para la estimación de salud de la prensa de peletizado provienen de 2 líneas de producción las cuales tienen características similares en cuanto a maquinaria, dieta y producción, provienen de un periodo de un año entre mayo de 2018 hasta mayo de 2019 que fueron entregados por la empresa WSP.

Los datos se pueden dividir en las siguientes 4 fuentes de información de las cuales provinieron:

- ***Sensores de planta:*** Estas características provienen del monitoreo de las diferentes variables de control que se necesitan en el proceso, comúnmente siendo corriente o potencias de alguna maquinaria, temperaturas en algún sub proceso u otras variables que representan el estado de distintas maquinas.
- ***Recetas e ingredientes:*** Estas características representan las diferentes recetas y la presencia de los ingredientes respectivos que están pasando por la línea. Las recetas dependerán de los pedidos que den los diferentes clientes y las necesidades nutritivas que tengan sus animales. Los ingredientes dependerán de las recetas y la disponibilidad que se tenga, lo cual puede implicar utilizar diferentes combinaciones de ingredientes para obtener el mismo perfil nutritivo de la receta original.
- ***Registros de mantención:*** Estas características provienen de los registros de fallas y mantenciones que se hacen en las diferentes líneas. Con estos registros se puede obtener los tiempos desde la última falla, fechas de la última mantención, razones de las fallas, tiempos de parada y reemplazos de piezas, pero también debido a que son datos ingresados por los operadores o encargados de mantención, pueden tener diferencias e inexactitudes debido al juicio de cada persona, siendo una fuente de incertidumbre.
- ***Contexto:*** Estas características no provienen explícitamente de alguna fuente de información específica y representas comúnmente características temporales como día de la semana, hora del día, etc.

### 3.2.1. Caracterización de modo de falla a modelar

Para lograr una caracterización del modo de falla se hizo un análisis de los distintos modos de fallas, analizando las causas raíces, consecuencias, frecuencia, riesgo de seguridad, componentes afectados y acciones de mantenimiento. Además, para el indicador de falla, se seleccionó el modo de falla que tuviera un efecto considerable en la producción de la planta y una frecuencia importante en comparación a las otras. Esto se logró analizando los registros de mantención y recopilando conocimiento experto de los operadores.

Con el proceso anterior se identifico un modo de falla que cumpliera con estos requisitos, en donde hubiera la mayor cantidad de eventos de fallas para obtener una mejor caracterización. Este modo de falla se le denomina en los documentos como "Equipo atollado", lo cual se caracteriza como una obstrucción de material en los agujeros del molde, por lo cual no deja un correcto flujo para la transformación de la masa en pellet, implicando un exceso de acumulación de material en la prensa y una detención no programada del proceso de pelletizado.

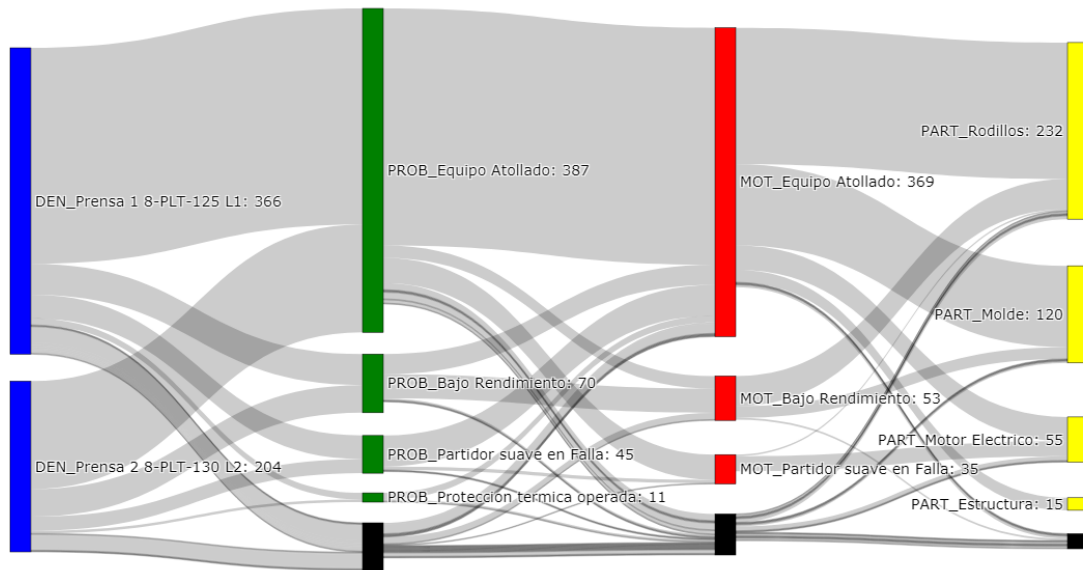


Figura 3.2: Diagrama Sankey de las fallas registradas de las prensas de pellet. Maquinaria (Azul), Problema (Verde), Motivo (Rojo) y parte en falla (Amarillo).

En la Figura 3.2 se puede observar un diagrama Sankey, en el cual con el ancho de las categorías y flujo entre ellas se puede observar la cantidad de fallas que caen las diferentes categorías, y podemos observar la gran cantidad de fallas que tiene como problema y motivo un "Equipo Atollado" además los cuales están muy relacionados con las partes del rodillo y el molde la prensa. En el rectángulo negro se contienen los elementos de menor porcentaje que provienen de la misma categoría que las divisiones verticales.

Este modo de falla además de estar muy relacionado con el funcionamiento y estado de las piezas internas de la prensa, también el conocimiento experto de los operaciones se sabe que existe una conexión con la calidad del proceso de acondicionamiento que tiene la mezcla y los ingredientes a utilizar, dado que esto define las propiedades físicas que tiene el producto antes del proceso de peletizado, lo cual puede dificultar el actuar de la prensa.

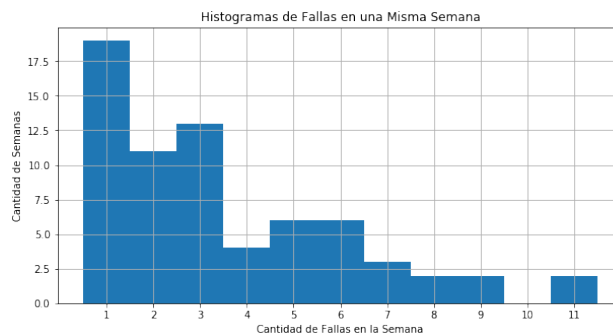


Figura 3.3: Histograma de tiempos de falla.

Para caracterizar de mejor forma el modo de falla de atolle se hizo un análisis de los tiempos de falla, lo que sería el tiempo entre fallas o el tiempo entre acciones de mantenimiento adecuadas y siguiente falla. En la Figura 3.3 podemos ver que en la mayoría de las semanas en donde suceden fallas, solo sucede una vez, aunque existe un porcentaje considerable de semanas que tienen más de una falla de atolle. Analizando con mas detalle se puede observar que más de la mitad de las semanas con fallas, tiene más de una falla de atolle.

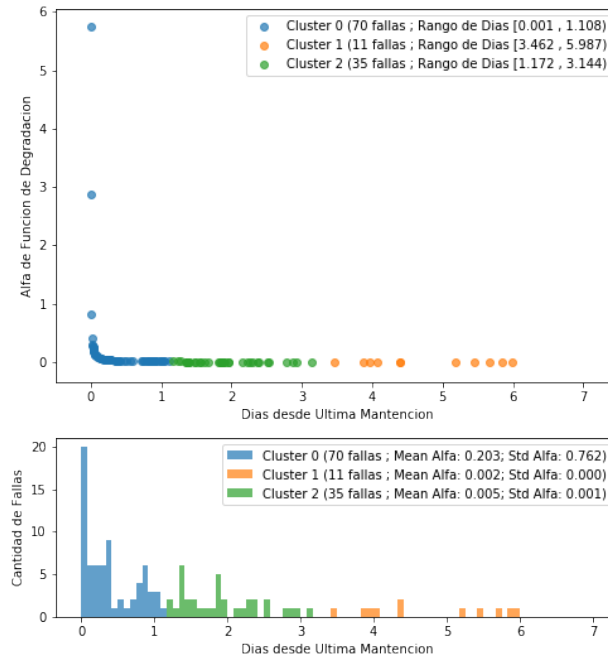


Figura 3.4: Cluster de los parámetro del modelo de falla y los tiempos entre fallas correspondiente.

Para filtrar los tipos de falla se hizo un clúster en función del parámetro utilizado en el modelo de degradación. En la Figura 3.4 se puede ver la caracterización de los diferentes clústeres que se lograron utilizando K-means y el tiempo entre fallas correspondiente.

Finalmente, al observar los clusters que se produjeron con K-means, la cantidad de eventos registrados y la varianza de los tiempos y parámetros del modelo, se tomo la decisión de dividir el dataset en dos,  $D_{<24h}$  y  $D_{>24h}$ , los cuales consistieron en datasets con las series de tiempos de fallas que contenían los datos de menos de 24 horas con 66 eventos y mayores a 24 horas con 50 eventos, respectivamente.

### 3.3. Pre-procesamiento de Datos

Para facilitar el aprendizaje del estimador es práctica común en redes neuronales normalizar los valores de las variables de entrada para acelerar el entrenamiento y mejorar la búsqueda de los

óptimos. Para esto se utilizó una normalización robusta.

El caso normal de la normalización consiste en que los valores de los datos de entradas se redefinen en función de la media y la desviación estándar de cada dato, según la ecuación 3.1.

$$\bar{x}_i = \frac{x_i - \mu_{x_i}}{\sigma_{x_i}} \quad (3.1)$$

Donde  $x_i$  es la característica  $i$  de los datos  $x$ ,  $\mu_{x_i}$  es la media de esa característica  $i$ ,  $\sigma_{x_i}$  es la desviación estándar de la característica  $i$  y  $\bar{x}_i$  es la característica  $i$  normalizada de los datos.

Pero en caso de la normalización robusta, se hace en función de la mediana y los rangos de los cuartiles, según la ecuación 3.2.

$$\bar{x}_i = \frac{x_i - \text{mediana}(x_i)}{RQ(x_i)} \quad (3.2)$$

Donde  $RQ$  es el rango intercuartil, que para este caso se uso el rango entre el cuartil 5 y 95, o sea la diferencia entre estos 2 cuartiles.

Se decidió utilizar una normalización robusta debido a que los datos de la maquinaria contienen mucho outlier y no todas se distribuyen de igual forma, provocando que los métodos de normalización comunes no se desempeñen correctamente.

Las creación de características fue a través de ventana de tiempos de corto tiempo corto (2 minutos) y con retrasos, y una ventana de largo tiempo (1 hora) como se ejemplifica en la Figura 3.5.

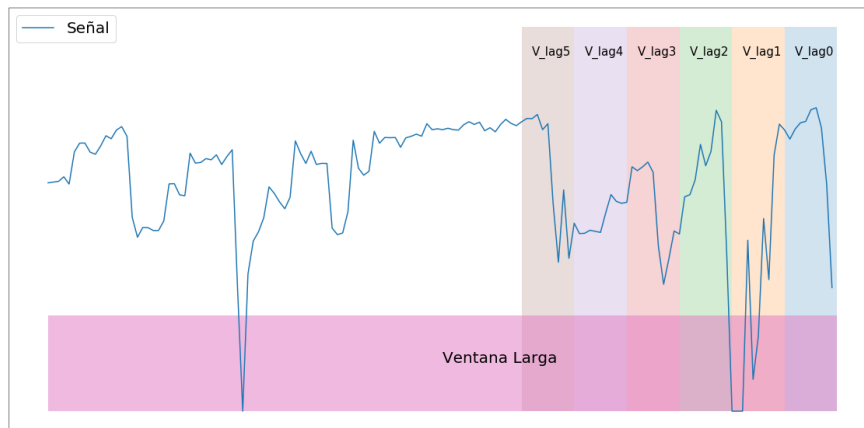


Figura 3.5: Ejemplo del uso de ventanas para creación características.

Con los datos de señales de las ventanas de tiempo se evaluaron las siguientes funciones:

- Media
- Desviación Estándar

- Kurtosis
- Asimetría
- Máximo
- Mínimo
- Integral (Trapezoidal)
- Perímetro

Los datos provenientes de diferentes fuentes de información y con diferentes tiempos de muestreos, se le extrajeron características utilizando ventanas de tiempo de corto y larga duración.

### 3.4. Modelo de Estimación de salud

Para el modelo de la estimación de salud de la máquina se tomó en cuenta el modo de falla específico que se desea trabajar, en este caso el atolle de la maquina debido a la formación de una capa de producto en el molde de la peletera. Para nuestro trabajo haremos la aproximación de la dinámica de la evolución de la falla, en el cual asumimos que el cambio en el porcentaje del molde obstruido es proporcional al mismo porcentaje, lo que implica que, al obstruirse el molde, lo demás se obstruye más rápido. Para la definición de la estimación de salud usaremos el área del molde obstruida ( $A_{obstruida}$ ) como la variable indicadora de salud, donde tendrá valores en el rango de 0 a 1 (0 representando un molde en óptimas condiciones y 1 un molde completamente obstruido, lo que provocaría una falla). La siguiente ecuación diferencial seria el modelo recién explicado:

$$\dot{A}_{obstruida} = k A_{obstruida} \quad (3.3)$$

donde se tiene que cumplir  $A_{obstruida}(0) \approx 0$  y  $A_{obstruida}(t_{falla}) = 1$ , lo que implica que la función de degradación es la siguiente:

$$\implies A_{obstruida}(t) = C e^{\frac{\ln(C)}{t_{falla}} t} \quad (3.4)$$

Donde  $t_{falla}$  es el tiempo en donde ocurre la falla y  $C$  debe ser un valor cercano a 0 para cumplir la condición de que la degradación en  $t = 0$  sea aproximadamente 0. Importante notar que un  $C$  más cercano a 0 implica un crecimiento mayor en la cercanía del tiempo de falla, y al contrario significa que una degradación más uniforme en el tiempo.

Esta función de degradación de la maquinaria se reiniciará a 0 después de cada falla o luego de una mantención rutinaria de la planta, por lo que estamos haciendo el supuesto que las acciones de mantención son efectivas y regresan a condiciones óptimas la prensa de peletizado.



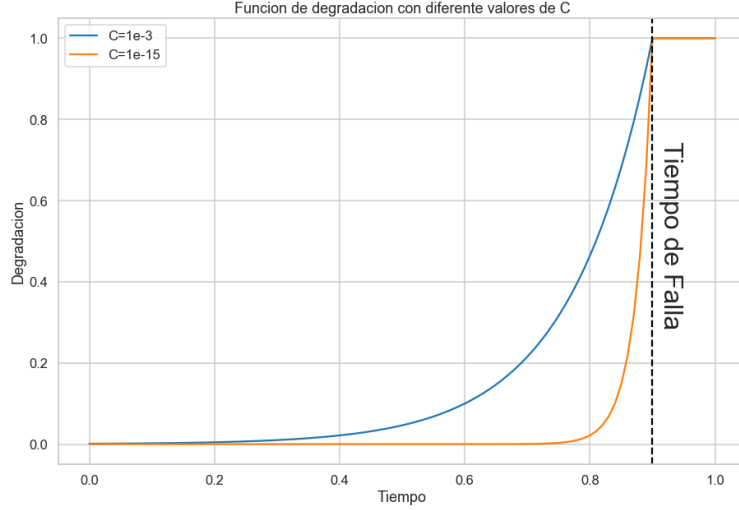


Figura 3.6: Función de degradación con diferentes valores de C.

Para el proceso de resultados en el dataset  $D_{>24h}$  se modelara la degradación con valores de C igual a  $1e-3$  y  $1e-15$ , esto con la intención de evaluar los resultados asumiendo una evolución lenta y otra rápida de degradación, mientras que para el caso del dataset  $D_{<24h}$  solo se evaluara con C igual a  $1e-3$ , esto debido a que la escala de tiempo de las fallas implicarían un evolución de degradación demasiado rápida para valores más altos de C.

### 3.5. Diseño del Estimador

Para la regresión del estado de degradación de la maquina se decidió utilizar una perceptron multicapas con un enfoque probabilístico, así entregando no solo la estimación, sino también la incertidumbre que posee.

Debido que los valores de la degradación están entre 0 y 1, se modeló la salida del modelo como una distribución Beta. Debido a eso, se utilizó como función de perdida el logaritmo de la verosimilitud para ajustar los valores de los parámetros de la distribución, la cual tendría la siguiente formula:

$$J(y, \bar{\alpha}, \bar{\beta}) = -\log(P_{beta}(y, \bar{\alpha}, \bar{\beta})) \quad (3.5)$$

$$J(y, \bar{\alpha}, \bar{\beta}) = -\log\left(\frac{(y)^{\bar{\alpha}-1}(1-y)^{\bar{\beta}-1}}{B(\bar{\alpha}, \bar{\beta})}\right) \quad (3.6)$$

$$J(y, \bar{\alpha}, \bar{\beta}) = -((\bar{\alpha} - 1)\log(y) + (\bar{\beta} - 1)\log(1 - y) - (\log(\Gamma(\bar{\alpha})) + \log(\Gamma(\bar{\beta})))) \quad (3.7)$$

Donde J es la función de costo, y es la etiqueta del dataset,  $\bar{\alpha}$  y  $\bar{\beta}$  son los parámetros estimados de la distribución beta y la salida del red neuronal y  $\Gamma$  es la función matemática Gamma.

Además, restringiendo los valores posibles de los parámetros  $\bar{\alpha}$  y  $\bar{\beta}$  podemos establecer ciertas

características que tendrá que tener la distribución. En específico los valores se restringieron a valores mayores a 1, ya que si no, la distribución se vuelve multimodal, lo cual no tendría sentido en el modelo del problema, y también se dejó como parámetro la cota menor del parámetro  $\alpha$ , esto debido que con esta cota podemos tender al estimador que favorezca la estimación de valores de degradación altos, ya que obliga que una estimación de valores bajos sea con mayor certeza que una estimación con valores altos, lo que para este caso es deseable debido a que las fallas son eventos muy esporádicos por lo que existen una menor cantidad de valores altos que de valores bajos, por ejemplo, acotando  $\alpha$  a 3 y  $\beta$  a 1, podemos comparar que para tener una distribución con moda en 0.75, los parámetros con mayor varianza serán  $(\alpha, \beta)$  igual a (3,1) con una varianza de 0.0375, mientras que para obtener una moda de 0.25 se tendrán los parámetros (3,9) con una varianza de 0.0114.

Para lograr la restricción de los valores de los parámetros, la salida de la red neuronal serán 2 valores que pasarán por la siguiente función:

$$F_{degradacion}(\vec{Y}_{nn}) = \begin{bmatrix} C_\alpha + \text{softplus}(Y_{nn}^1) \\ C_\beta + \text{softplus}(Y_{nn}^2) \end{bmatrix} = \begin{bmatrix} \bar{\alpha} \\ \bar{\beta} \end{bmatrix} \quad (3.8)$$

Donde  $Y_{nn}$  es la salida de la red después de las capas fully connected,  $(C_\alpha, C_\beta)$  son las cotas de los parámetros  $\alpha$  y  $\beta$ , y  $(\bar{\alpha}, \bar{\beta})$  son los parámetros de la distribución de salida. Como la función softplus tiene un rango de solo los  $\mathbb{R}^+$ , esto obliga que la salida de la función estará acotada por  $(C_\alpha, C_\beta)$ .

Para la optimización de los hiper-parámetros se utilizó una búsqueda de grilla para encontrar los valores óptimos de arquitectura y entrenamiento.

Para la selección de características se utilizó el método de permutación, el cual consiste en seleccionar una característica y permutarla dentro del dataset, luego este nuevo dataset es evaluado y se ve la diferencia en el desempeño de la red utilizando la función de costo, si es que la función de costo aumento, significa que la variable era importante, si es que la función de costo disminuye o se mantuvo igual, significa que la variable no aportaba ninguna información relevante a la decisión del modelo. Esto se repite para cada característica y luego iterativamente se van eliminando algún numero o porcentaje de características que tengan la mayor disminución en el cambio de desempeño.

## Capítulo 4

# Análisis de Resultados

### 4.1. Parámetros Estimator

La arquitectura y parámetros óptimos para cada sub conjunto de datos fueron las siguientes:

C	Dataset		
	$D_{<24h}$	$D_{>24h}$	
	<b>1e-3</b>	<b>1e-3</b>	<b>1e-15</b>
Arquitectura	[256,128,64,2]	[256,128,64,2]	[128,64,32,2]
Factor de Regularización (L1)	5e-8	5e-8	5e-9
Dropout	0.2	0.3	0.1
$C_\alpha$	3	3	3
$C_\beta$	1	1	1

Tabla 4.1: Parámetros del Modelo.

Importante notar que al entrenar las redes, los resultados variaban bastante y en muchos casos se convergía a una solución constante o muy cercana a 0.

### 4.2. Filtrado de Característica

Al utilizar el método de permutación para filtrar las características de entrada de la red, se dio que al ordenar las que tenían un mayor impacto en la función de costo se repetían en los diferentes dataset, aunque en diferentes magnitudes.

La media de la temperatura de salida de la peletera y de la boquilla del FEX aparecen consistentemente entre las que tenían un mayor impacto en la función de costo, con diferentes retrasaos de las ventanas de tiempo. Esto hace sentido con la documentación encontrada del proceso de peletizado en donde la temperatura tiene un factor importante en la calidad del pellet.

El tiempo transcurrido desde la ultima falla también es una característica importante, lo que se puede percibir notoriamente en los resultado de la estimación, en donde al aumentar el tiempo, la degradación aumenta, lo cual es consistente con el modelo que se definió.

La última característica que tenía un efecto importante en la función de costo al utilizar el método de permutación, fue la apertura del lineator, la cual representa de cierta forma la distancia entre los rodillos de la prensa. En el caso de degradación rápida fue donde tuvo la mayor variación de la función de costo.

### 4.3. Estimación de degradación

A continuación, al utilizar los parámetros de la sección anterior, se presentan los resultados de la red neuronal. Para la evaluación de la estimación se utilizó el MSE (Mean Squared Error), en la Tabla 4.2 están los resultados de los diferentes dataset y diferentes modelos de degradación.

C	Dataset		
	$D_{<24h}$	$D_{>24h}$	
	<b>1e-3</b>	<b>1e-3</b>	<b>1e-15</b>
MSE Total	0.0683897	0.0870862	0.0224929
MSE Promedio por Serie	0.0297921	0.0139947	0.0064029

Tabla 4.2: Resultados de los Modelos.

La métrica MSE total consiste en el error cuadrático medio de todo los datos de las series de tiempo del conjunto de test, mientras que la métrica MSE promedio por serie, es el promedio del error cuadrático medio de las diferentes series de tiempo del conjunto de test.

Al observar los errores de los diferentes modelos, se puede notar que el valor de C tiene un impacto en la escala de los errores, esto no es necesariamente debido a que se tiene un indicador que explique de mejor forma la condición de la maquina, sino que es debido a que al tener un valor C menor, también se obtiene un mayor porcentaje de valores cercanos a 0 lo que implica que el estimador puede sesgarse hacia valores menores y en promedio tener errores menores.

#### 4.3.1. Dataset $D_{>24h}$

El dataset  $D_{>24h}$  fue evaluado con dos modelos de degradación, uno lento, enfocado más en un desgaste debido al uso de la máquina y sus componentes, y uno rápido, enfocado más en la operación de la maquinaria.

Los resultados del modelo con C igual a 1e-3 se obtuvieron los mejores resultados, en donde la gran mayoría de las series de tiempos obtuvieron una estimación de degradación cercana a la del modelo.

En los resultados se pudo observar que el rendimiento del indicador en caso donde no esta dentro de los limites del rango del tiempo de falla del dataset (1 a 7 días), se obtiene buenos resultados

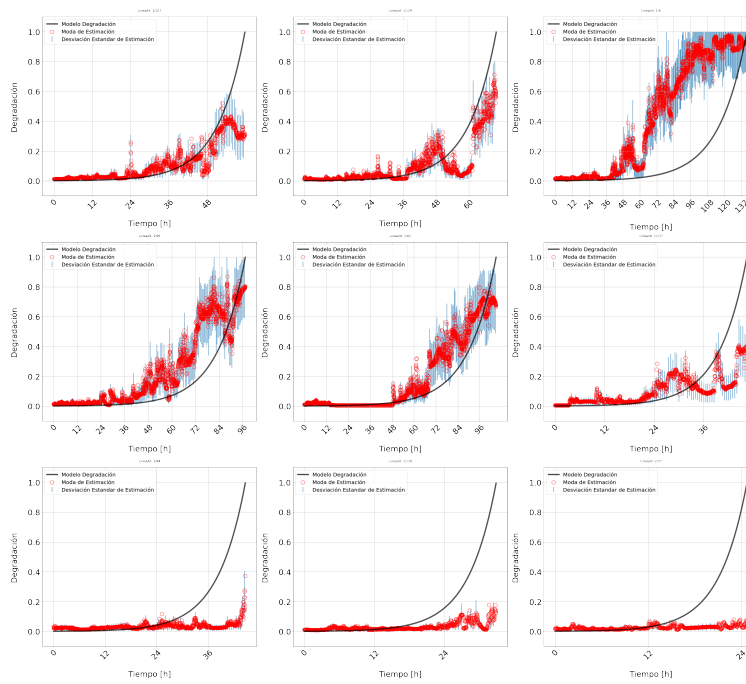


Figura 4.1: Resultados de modelo de degradación lenta en dataset  $D_{>24h}$ . Modelo de degradación (línea negra) y Estimación de red neuronal (círculos rojos).

como se puede observar en la Figura 4.2 la diferencia de los rangos de tiempos entre las dos series de tiempos es importante, ambas estimaciones obtuvieron resultados similares.

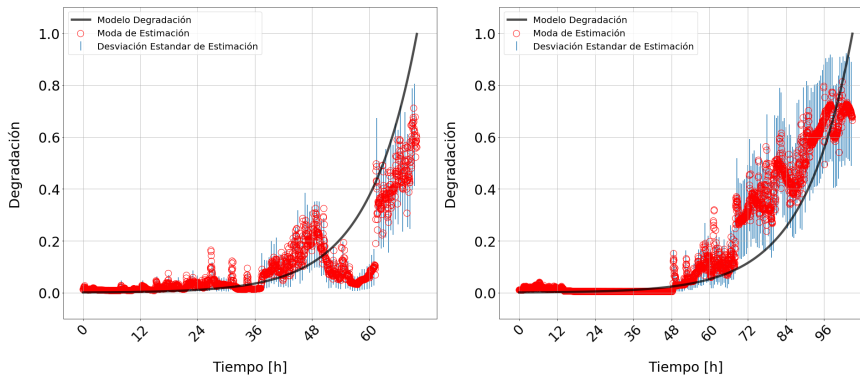


Figura 4.2: Comparación resultados de modelo lento de degradación en dataset  $D_{>24h}$ . Modelo de degradación (Negro) y Estimación de red neuronal (Rojo).

Pero en los límites de los rangos de tiempos del dataset el indicador tiende a tener un menor desempeño, como se nota en la Figura 4.3, el estimador tiende a devolver valores cercanos a 0 (para fallas con tiempo cercanos a 1 día) o saturarse en 1 (para fallas con tiempos cercanos a la semana).

En cuanto a la incertidumbre del indicador, podemos ver que con el uso de la distribución beta y la cotas de los parámetros, se obtuvo el comportamiento deseado en donde para valores cercano

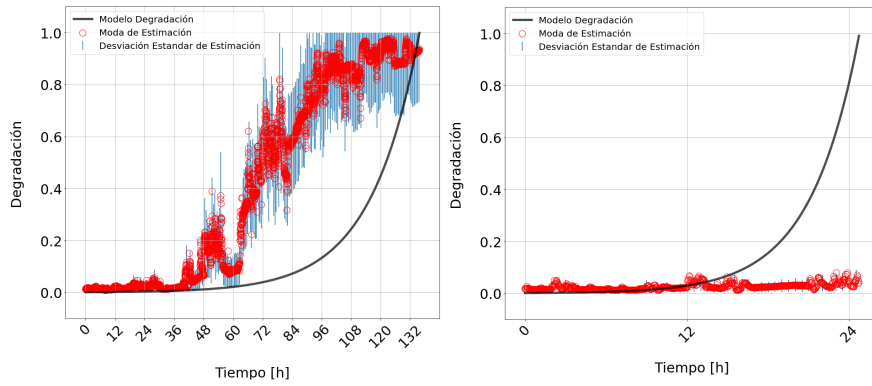


Figura 4.3: Comparación resultados de modelo de degradación lenta en dataset  $D_{>24h}$  en tiempos limites. Modelo de degradación (Negro) y Estimación de red neuronal (Rojo).

a 0, la incertidumbre es menor en comparación que para los valores de degradación mas altos.

En cuanto a los resultados del modelo con C igual a  $1e-15$  se obtienen mucha más varianza en los resultados, en donde el índice entrega una mayor cantidad de "falsos positivos", es decir incremento de valores de degradación no cercanos al tiempo de falla.

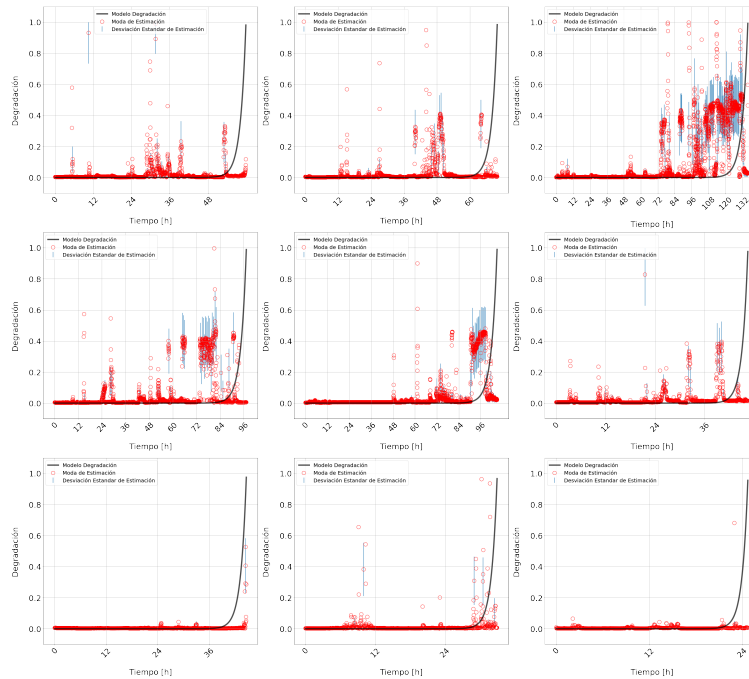


Figura 4.4: Resultados de modelo rápida de degradación en dataset  $D_{>24h}$ . Modelo de degradación (línea negra) y Estimación de red neuronal (círculos rojos).

En la Figura 4.4 podemos ver como en diferentes series de tiempos existen estos falsos positivos distribuidos en diferentes partes de la serie de tiempo, se puede hipotetizar que esto implica que existen varios estados cercanos o similares a fallas, pero que no progresaron más y son esos estados

los que están siendo identificado por el estimador. Lo anterior podría ser válido si es que también se reflejara en un buen desempeño del estimador en la cercanía del tiempo de falla, lo cual no sucede en la mayoría de las series de tiempo, por lo que podríamos argumentar que este modelo de degradación acelerada no explica correctamente la dinámica del modo de falla y estado de salud de la máquina.

#### 4.3.2. Dataset $D_{<24h}$

El dataset  $D_{<24h}$  fue evaluado solo con el modelo de degradación lento debido a los rangos de tiempos con los que se trabajan.

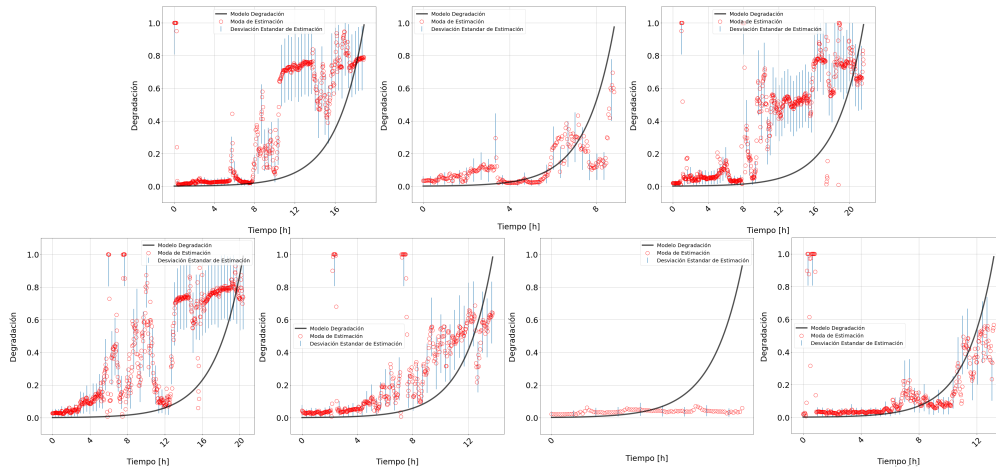


Figura 4.5: Resultados de modelo de degradación lenta en dataset  $D_{<24h}$ . Modelo de degradación (Negro) y Estimación de red neuronal (Rojo).

En estos resultados de la Figura 4.5 podemos notar que sucede algo similar que en el dataset  $D_{>24h}$ , en donde para tiempo de falla muy bajos, el estimador no identifica una degradación en el proceso. Pero con excepción a esos casos, los resultados tienden a obtener valores altos en la cercanía de los tiempos de falla, además que los casos en donde existen tiempos de fallas tan cortos, muy posiblemente se deba a acciones de mantenimiento imperfectas.

Al observar podemos ver que existen casos en donde la evolución de la estimación es más súbita, lo cual puede implicar que debido a que este dataset posee rangos de tiempo más pequeño, se pueda detectar condiciones de posible mayor degradación, en vez de ajustarse a una degradación más paulatina.

## Capítulo 5

# Conclusiones

Hoy en día la industria en la búsqueda de optimizar sus procesos y mejorar la toma de decisiones, se ha empezado a tomar mayor consciencia de los datos que poseen y buscar crear un mayor valor de ellos. En conjunto con el aumento en la capacidad de almacenaje de datos y el crecimiento en investigación sobre métodos de inteligencia computacional, las implementaciones de estas técnicas se han vuelto más comunes, siendo aplicadas en diversas áreas industriales y comerciales.

En esta Memoria de Título se propuso la implementación de un indicador para el pronóstico de un modo de falla en una planta de alimentos, en específico el atolle de una prensa peletizadora. En nuestro diseño se utilizó una red neuronal en donde se utilizaron los datos de diferentes sensores, productos que se utilizaban y registros de mantención para obtener el indicador y también la incertidumbre de la predicción. Además, se hizo con el enfoque de ser utilizado para la predicción de tiempos de falla utilizando un enfoque bayesiano para la actualización de un modelo a priori.

Lo que se ha podido concluir con los resultados es que existe la oportunidad de poder implementar un sistema de pronóstico de fallas utilizando el indicador, en específico el modelo de degradación lenta fue el que obtuvo mejores resultados en ambos dataset, en donde la mayoría de los casos se obtenían resultados similares a la tendencia de la degradación de la máquina, además es importante notar que por la disminución de los datos debido al filtrado de fallas y división del dataset, el aporte que trae la implementación de un modelo de degradación entrega una mayor información que puede compensar esa falta de datos y un enfoque puramente basado en datos. Pero es necesario revisar diferentes aspectos que necesitan una mayor atención para la implementación de un sistema de predicción de fallas en un ámbito de producción.

Un aspecto que provocó dificultades fue la convergencia del modelo. Debido al tipo de función del modelo de degradación, el largo variable de las series de tiempos y el uso de un acercamiento probabilístico al problema, la red neuronal tendía a entregar valores bajos, provocando que muchos modelos convergerán a la media de la degradación de cada serie de tiempo. Esto se pudo resolver



con el ajuste de los hiperparámetros, principalmente de las cotas inferiores de los parámetros de la distribución beta, pero para obtener una mejor fiabilidad del sistema será necesario un mayor volumen de datos, no solo para el entrenamiento, si no también para su debida evaluación.

Otro aspecto que es necesario revisar y se pudo vislumbrar a través de la metodología utilizada fue la relevancia que tienen ciertas variables para el modelo, principalmente la temperatura de los procesos, en donde se diviso la importancia de esa variable para el desempeño del indicador, por lo que será necesario hacer un estudio más preciso para ver si a través de esta información se puede mejorar el modelo de degradación o la misma red neuronal.

Para un trabajo futuro se deberá enfocar en 2 ámbitos, en la calidad de los datos y la modelación del problema. En cuanto a la calidad de los datos se debería trabajar en una propuesta para la identificación más específica del tipo de falla para el equipo de mantención de la planta, logrando una mejor caracterización de las fallas y las causas raíces, esto debido que en este trabajo se utilizaron los registros de mantención, pero muchas veces en los registros existían campos que eran muy similares y creaban una incertidumbre sobre el modo falla que había ocurrido. También proponer otras variables de interés de los procesos, lo cual puede incluir la implementación de nuevo sensores o el registro de variables que se utilizan solamente para control y que no se almacenan. Y en cuanto a la modelación del problema, se propone que al mejorar la calidad y tamaño del dataset, se podría probar la implementación de modelos más complejos, pero con mayor enfoque a series de tiempo, como RNN o LSTM.

# Bibliografía

- [1] C. Chen, B. Zhang, and G. Vachtsevanos. Prediction of machine health condition using neuro-fuzzy and bayesian algorithms. *IEEE Transactions on Instrumentation and Measurement*, 61(2):297–306, Feb 2012.
- [2] CPM. The pelleting process. <https://www.cpm.net/downloads/Animal%20Feed%20Pelleting.pdf>. Accessed: 2019-12-09.
- [3] M. Daigle and C. S. Kulkarni. A battery health monitoring framework for planetary rovers. In *2014 IEEE Aerospace Conference*, pages 1–9, March 2014.
- [4] Y. Gal. Uncertainty in deep learning. pages 127–131, 2016.
- [5] M. R. A. H. George Vachtsevanos, Frank Lewis and B. Wu. *Intelligent Fault Diagnosis and Prognosis for Engineering Systems*. John Wiley & Sons, Inc., Hoboken, New Jersey, 2006.
- [6] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016. <http://www.deeplearningbook.org>.
- [7] L. He. Application of bayesian network in power grid fault diagnosis. In *2008 Fourth International Conference on Natural Computation*, volume 1, pages 61–64, Oct 2008.
- [8] J. Luo, K. R. Pattipati, L. Qiao, and S. Chigusa. Model-based prognostic techniques applied to a suspension system. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 38(5):1156–1168, Sep. 2008.
- [9] M. Orchard, G. Kacprzynski, K. Goebel, B. Saha, and G. Vachtsevanos. Advances in uncertainty representation and management for particle filtering applied to prognostics. In *2008 International Conference on Prognostics and Health Management*, pages 1–6, Oct 2008.
- [10] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Neurocomputing: Foundations of research. chapter Learning Representations by Back-propagating Errors, pages 696–699. MIT Press, Cambridge, MA, USA, 1988.

- [11] A. Saxena, J. Celaya, E. Balaban, K. Goebel, B. Saha, S. Saha, and M. Schwabacher. Metrics for evaluating performance of prognostic techniques. In *2008 International Conference on Prognostics and Health Management*, pages 1–17, Oct 2008.
- [12] A. S. Yoon, T. Lee, Y. Lim, D. Jung, P. Kang, D. Kim, K. Park, and Y. Choi. Semi-supervised learning with deep generative models for asset failure prediction. *CoRR*, abs/1709.00845, 2017.
- [13] G. Yu and L. Zhang. An intelligent approach for bearing fault diagnosis based on bayesian networks and alpha-stable distribution. In *2015 International Conference on Computational Intelligence and Communication Networks (CICN)*, pages 76–78, Dec 2015.
- [14] B. Zhang, C. Sconyers, C. Byington, R. Patrick, M. E. Orchard, and G. Vachtsevanos. A probabilistic fault detection approach: Application to bearing fault detection. *IEEE Transactions on Industrial Electronics*, 58(5):2011–2018, May 2011.
- [15] W. Zhang, C. Li, G. Peng, Y. Chen, and Z. Zhang. A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load. *Mechanical Systems and Signal Processing*, 100:439–453, 02 2018.