

PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO
FACULTAD DE INGENIERÍA
ESCUELA DE INGENIERÍA INFORMÁTICA

**DIAGNÓSTICO DE FALLA DE RODAMIENTO
UTILIZANDO DESCOMPOSICIÓN MODAL
EMPÍRICA Y DEEP LEARNING**

GERD PATRICIO KIMER ROZAS

Profesor guía: **Nibaldo Rodríguez Agurto**

Profesor co-referente: **Guillermo Cabrera Guerrero**

DICIEMBRE, 2017

Dedicatoria

A mis padres quienes con todo su esfuerzo, dedicación y amor me brindaron la oportunidad de estudiar y las herramientas para llegar a ser quien soy. A mi novia por apoyarme incondicionalmente en los momentos difíciles, celebrar mis momentos felices y ser mi compañera de vida. A mis hermanos maternos por estar desde siempre a mi lado dándome fuerzas, amor y comprensión en este largo camino.

Agradecimientos

A mis amigos por su apoyo, ayuda e infinito aguante en estos 6 años. A mi profesor guía por entregarme sus conocimientos, ayuda y enseñanzas. A todos los que en este camino me han enseñado y formado como ingeniero.

Resumen

En la industria, para evitar pérdidas monetarias, es importante conocer el estado de los rodamientos de las máquinas dado que representan el 40 % del total de las averías. Una manera de conocer si los rodamientos tienen fallas es a través de modelos matemáticos que analicen las vibraciones de la máquina en funcionamiento. Estos modelos se componen de dos etapas, la primera consiste en extraer características desde la vibración y la segunda consiste en utilizar un clasificador para identificar la falla. Esta investigación propone un diagnosticador de fallas que extrae características por medio de la descomposición modal empírica y clasifica la falla gracias al aprendizaje profundo. El modelo fue probado con datos del CWRU buscando diagnosticar 10 tipos de fallas de rodamiento bajo condiciones de operación variable. Los resultados obtenidos muestran que el modelo logra una exactitud promedio del 97.23 %, donde su fortaleza se evidencia en los fallos en etapas iniciales. Finalmente los resultados demuestran que el modelo es una excelente técnica para realizar el mantenimiento predictivo de rodamientos en maquinarias industriales.

Palabras clave: *descomposición modal empírica, aprendizaje profundo, autoencoder sparse, diagnosticador de falla de rodamientos.*

Abstract

In the industry, to avoid monetary losses, it is important to know the state of the bearings of the machines since they represent 40 % of the total of the breakdowns. One way to know if the bearings have failures is through mathematical models that analyze the vibrations of the machine in operation. These models are composed of two stages, the first consists of extracting characteristics from the vibration and the second consists of using a classifier to identify the fault. This research proposes a fault diagnosis that extracts characteristics through empirical modal decomposition and classifies the failure thanks to deep learning. The model was tested with data from CWRU seeking to diagnose 10 types of bearing failures under varying operating conditions. The results obtained show that the model achieves an average accuracy of 97.23 %, where its strength is evident in the failures in early stages. Finally, the results show that the model is an excellent technique to carry out the predictive maintenance of bearings in industrial machineries.

Keywords: *empirical mode decomposition, deep learning, autoencoder sparse, bearing fault diagnosis.*

Índice

Lista de Figuras	VII
Lista de Tablas	VIII
1. Introducción	1
2. Definición de objetivos	4
2.1. Objetivo General	4
2.2. Objetivos Específicos	4
3. Marco teórico	5
3.1. Estado del arte	5
3.2. Descomposición modal Empírica (EMD)	7
3.3. Descomposición modal Empírica Ensamblada (EEMD)	9
3.4. Deep Learning	10
3.4.1. Autoencoder	11
3.4.1.1. Autoencoder Sparse	12
3.5. Clasificador Softmax	14
4. Modelo de diagnóstico de fallas propuesto	15
4.1. Pre Proceso	15
4.2. Extracción de Características	15
4.2.1. Características utilizadas	16
4.3. Clasificador Inteligente	16
4.4. Métricas utilizadas	17
5. Presentación de resultados	19
5.1. Descripción del experimento utilizado	19
5.2. Conjunto de datos utilizados	20
5.3. Ajustes del modelo propuesto	21
5.3.1. Pre procesamiento	21
5.3.2. Extracción de características	22
5.3.3. Aprendizaje profundo	24
5.4. Primer Experimento	27
5.4.1. Extracción de características con EMD	28
5.4.2. Extracción de características con EEMD	29
5.5. Segundo Experimento	30
5.6. Discusión de resultados	32

6. Conclusiones	34
7. Referencias	36

Lista de Abreviaturas o siglas

1. EMD: Descomposición modal empírica
2. EEMD: Descomposición modal empírica ensamblada
3. IMF: Función modal intrínseca
4. ANN: Red neuronal artificial
5. DNN: Red neuronal profunda
6. KF: Factor de curtosis
7. CR: Factor de cresta
8. RMS: Raíz de la media al cuadrado
9. SVM: Máquina de vector de soporte
10. ELM: Máquina de aprendizaje extremo
11. SLNN: Red neuronal de una capa
12. CWRU: Case Western Reserve University

Lista de Figuras

1.	Arquitectura de un Auto Encoder[18]	11
2.	Diagrama de bloques modelo propuesto	15
3.	Foto de la configuración utilizada por el CWRU[24]	19
4.	Señales de Vibración	21
5.	EMD diferentes señales	23
6.	EEMD diferentes señales	24
7.	Exactitud vs Nodos ocultos	25
8.	Red Neuronal Profunda utilizada	25
9.	Proporción de Sparsity vs Exactitud	26
10.	Coficiente beta vs Exactitud	26
11.	Coficiente Lambda vs Exactitud	27
12.	EMD - Exactitud en 40 intentos	28
13.	EMD - F-score promedio por clases en 40 intentos	29
14.	EEMD - Exactitud en 40 intentos	29
15.	EEMD - F-score promedio por clases en 40 intentos	30
16.	Red Neuronal 1 AE utilizada	31
17.	Exactitud en 40 intentos	31
18.	F-score promedio por clases en 40 intentos	32

Lista de Tablas

1.	Algoritmo EMD	9
2.	Algoritmo EEMD	10
3.	Algoritmo Propuesto	17
4.	Conjunto de datos utilizado	20
5.	Clasificación de cada tipo de falla	20
6.	Exactitud en diversos tamaños de segmentación	21
7.	Tamaño del conjunto de datos y sus exactitudes	22
8.	Exactitud bajo distintos ruidos blanco	23
9.	Comparación distintas configuraciones	32

1. Introducción

La industria, por medio de una secuencia de esfuerzos, crea desde materias primas bienes que son comercializados para generar ingresos y volver a comenzar el ciclo productivo. En este proceso de creación intervienen tanto humanos como máquinas y, gracias al trabajo en conjunto es posible obtener el producto final. Las maquinarias, para cumplir sus funciones, se componen de una serie de piezas mecánicas y elementos rotativos genéricos, como por ejemplo: rodamientos, engranajes, rotores, entre otros. Estos elementos tarde o temprano sufrirán un desperfecto que descompondrá la máquina y como consecuencia generarán una detención de la línea productiva y en efecto pérdidas monetarias, incluso, pérdidas humanas en determinadas situaciones [1].

Con el fin de reducir el riesgo de una paralización en la línea productiva debido a un fallo inesperado, los ingenieros encargados del mantenimiento deben fijar detenciones periódicas preventivas. A pesar de contar con planes de mantenimiento basados en la confiabilidad, estos no aseguran la existencia de una falla en los componentes generando una mala optimización del tiempo en funcionamiento de la maquinaria. En este sentido existe la necesidad de contar con una herramienta automatizada que determine la existencia y severidad de un fallo sin la necesidad de detener la máquina. Los modelos ingenieriles que entregan solución a esta problemática, permiten asegurar la existencia de una falla al momento de detener una máquina y así optimizar el tiempo en funcionamiento y evitar pérdidas monetarias.

En relación con los elementos rotativos que forman parte de una máquina, es de suma importancia fijar la atención en los rodamientos, dado que aproximadamente el 40% de los desperfectos son a causa de estos[2]. Los rodamientos se componen de 4 elementos: anillo exterior, anillo interior, elementos rodantes y jaula; cada una de estas piezas sufre desperfectos a causa de factores tanto internos como externos. Por una parte, dentro de los factores internos se encuentra el estrés debido a vibraciones, la excentricidad del anillo interno y la corriente del rodamiento. Por otra, en los factores externos destacan la contaminación y corrosión, falta de lubricación, mal montaje del rodamiento y exceso de carga. En este sentido existen principalmente dos tipos de fallas en los rodamientos: falla de un solo punto y el deterioro de la superficie de apoyo. En consecuencia, estos desperfectos producen un aumento en las vibraciones, además del nivel de ruido generado por el rodamiento cuando se encuentra en funcionamiento[3].

Actualmente, gracias a la utilización de sensores en las máquinas, como los acelerómetros encargados de medir la oscilación de un cuerpo, se ob-

tienen señales de vibración que describen el comportamiento dinámico de los rodamientos en funcionamiento. Estas ondas tienen la particularidad de ser no lineales, no estacionarias y multicomponente. Esto quiere decir que, en una señal de esta índole, el diagnóstico de fallas de rodamientos es dificultoso de realizar y requiere de un método sofisticado para lograr buenos resultados[3].

En general, los modelos de diagnóstico de fallas que analizan las señales de vibración, con el fin de determinar si el elemento analizado presenta un desperfecto, se componen de dos etapas. La primera etapa consiste en realizar una extracción o selección de características desde la señal y, la segunda etapa, se basa en utilizar un clasificador inteligente que determine la severidad de la falla. En la bibliografía especializada, los modelos de diagnóstico en la primera etapa utilizan técnicas del dominio del tiempo, frecuencia y tiempo-frecuencia, con el propósito de obtener características de la onda en estudio.

Las técnicas para extraer características en el dominio del tiempo incluyen métodos estadísticos por ejemplo: la desviación estándar, skewness, factor de kurtosis, factor de cresta, raíz de la media al cuadrado, entre otros [4]. Para obtener características en el dominio de la frecuencia el método más utilizado la transformada rápida de Fourier, que proporciona un espectro de frecuencia para la señal de vibración [5]. En el área del análisis tiempo-frecuencia los métodos más relevantes son la descomposición modal empírica[6] y ensamblada [7], la transformada wavelet [8] y la transformada de Hilbert.

Seguido de la etapa de extracción de características, los modelos de diagnóstico de fallas -en su mayoría- utilizan sistemas, tanto lineales como no lineales, que identifican y clasifican automáticamente el estado del rodamiento. Esta labor la realizan algoritmos que toman las características obtenidas y las utilizan para encontrar patrones que les permitan distinguir de forma correcta un rodamiento con fallas. Actualmente, dentro de los algoritmos que entregan mejores resultados en la detección de fallas destacan las redes neuronales artificiales [9], máquinas vectoriales de soporte[10], redes neuronales profundas[11], entre otras.

En esta investigación, se presenta un modelo de detección de fallas en rodamientos que utiliza el método de análisis de señales de la descomposición modal empírica y un diagnosticador basado en el aprendizaje profundo. La descomposición modal empírica se utiliza con el propósito de extraer las oscilaciones incrustadas en la señal de vibración del rodamiento que no son identificables por los métodos tradicionales de análisis de señales. En este sentido, las oscilaciones encontradas en la señal de vibración serán utilizadas

para realizar una extracción características por medio de métodos estadísticos. Luego, estas se utilizan de entradas al modelo de aprendizaje profundo con el fin de encontrar patrones que sean útiles al momento de detectar y clasificar desperfectos presentes en los rodamientos. Finalmente, el modelo de detección se compone de una etapa de pre entrenamiento que utiliza autoencoder sparse y una etapa de clasificación de fallas a través una red neuronal profunda.

2. Definición de objetivos

2.1. Objetivo General

Desarrollar un diagnosticador de severidad de fallos de rodamientos bajo velocidades variables utilizando descomposición modal empírica original y ensamblada y Deep Learning.

2.2. Objetivos Específicos

1. Descomponer señales de vibración con descomposición modal empírica original
2. Descomponer señales de vibración con descomposición modal empírica ensamblada.
3. Abordar la extracción de características desde las funciones modales intrínsecas utilizando características estadísticas.
4. Implementar un algoritmo de Deep Learning con Autoencoder Sparse para diagnosticar fallos 10 clases de fallo.
5. Ajustar los parámetros del modelo de aprendizaje profundo.
6. Comparar el rendimiento de extracción de características de descomposición modal empírica original versus ensamblada.
7. Comparar el rendimiento del algoritmo de Deep Learning frente a diferentes configuraciones.

3. Marco teórico

El presente apartado, expone, en primer lugar, el estado actual del área de diagnóstico de fallas en rodamientos con base en otras investigaciones, como estas han abordado la problemática y los resultados obtenidos. En segundo lugar, se detalla todo el sustento teórico que la presente investigación toma como referente para crear un modelo de diagnóstico de fallas de rodamientos. Finalmente, se explica la novedosa metodología a utilizar dentro de esta investigación y cuál es su objetivo dentro del diagnóstico de fallas de rodamientos.

3.1. Estado del arte

El análisis de las vibraciones emitidas por un rodamiento cuando se encuentra en funcionamiento ha sido una de las áreas más relevantes a la hora de diagnosticar la ubicación y severidad de fallas. Para realizar el análisis de estas señales existen diversos métodos, los que son categorizados según el dominio que utilizan; entre estas técnicas están el dominio del tiempo, de la frecuencia y tiempo-frecuencia. Luego de obtener características de las vibraciones se utilizan sistemas de aprendizaje automático que encuentran patrones para determinar las fallas presentes en un rodamiento.

Con base en las cualidades del fenómeno estudiado, la descomposición modal empírica ha demostrado ser una de las técnicas del análisis tiempo-frecuencia más poderosas y una de las más utilizadas en la literatura especializada. Gracias a esta técnica se logra obtener desde una señal las funciones modales intrínsecas que la componen, que expresan las oscilaciones mono modales incrustadas en una señal. Cabe destacar que al realizar una extracción de características sobre las funciones modales intrínsecas se obtiene una mayor información sobre el fenómeno estudiado.

En la literatura especializada se encuentra amplia variedad de ejemplos del cálculo de características sobre las funciones modales intrínsecas para describir el fenómeno de las fallas. Este es el caso de la investigación de Yang Yu [6] donde se realiza el cálculo de la entropía a la energía presente en las IMF. Esta técnica permite a través de una red neuronal identificar las fallas presentes en el anillo externo, interno o un estado normal. Otro ejemplo de este tipo de técnica es la investigación de Xiaoyuan Zhang [7] al igual que Yang Yu obtiene la entropía de la energía de las IMF y en base a ellas determina si existe la falla. Luego de saber que existe una falla realiza el cálculo de los valores singulares de las IMF para determinar de forma exacta cuál es la falla, su severidad y localización. Cuando el diagnóstico de

fallas en rodamientos es abordado desde la extracción de características toma gran importancia seleccionar características que representen verazmente el fenómeno en estudio. De esta forma la utilización de las características adecuadas nos entregará un diagnóstico certero sobre la existencia de una avería en el rodamiento. Al respecto, existen investigaciones como la de Jacek Dybała [12] donde gracias al cálculo de la curtosis en las funciones modales intrínsecas que representan el ruido en la señal logran determinar si un rodamiento presenta una falla. Este estudio realiza el cálculo de la curtosis a las señales del rodamiento saludable y con presencia de fallas, para luego con base en el valor encontrado discriminar la existencia de un desperfecto. Sin embargo, para lograr una observación con detalle, particularmente conocer la severidad o localización, se requiere mayor información de la señal y un algoritmo que aprenda los patrones sobre el fenómeno.

En la investigación de Wentao Mao [13] se utiliza el valor de la energía presente en las primeras cuatro IMF, como características para determinar fallas en rodamientos. Luego de obtener estas características se utilizan dos clasificadores, una máquina de vector de soporte y una máquina de aprendizaje extremo. Estos clasificadores al intentar identificar la severidad y localización de la falla, logran una exactitud del 76.96 % y del 66.95 % para la SVM y ELM respectivamente. Mientras que al intentar determinar solamente la severidad de la falla, la exactitud aumenta a valores entre un 84 % a 97.67 % y 82.76 % a 93.96 %, para la SVM y la ELM. En definitiva, para lograr un diagnóstico certero se debe crear un modelo que combine una correcta extracción de características sumado a un algoritmo de aprendizaje que amplifique los patrones de característicos de la falla.

En relación con los algoritmos de aprendizaje, existe un área relativamente nueva llamada aprendizaje profundo o Deep Learning. Estas técnicas permiten a través de grandes volúmenes de datos encontrar patrones en altos niveles de abstracción, los que permiten realizar complejas clasificaciones. Actualmente, existen diversas áreas donde el aprendizaje profundo ha realizado grandes avances, entre los que destacan el reconocimiento de imágenes, voz y escritura, la conducción autónoma, la automatización en la industria entre otras. La investigación de Ran Zhang [11] utiliza el aprendizaje profundo para determinar la existencia y localización de la avería tomando como entrada la señal en bruto del rodamiento. En esta investigación realizan una segmentación de la señal de vibración del rodamiento para luego, con cada segmento alimentar una red neuronal profunda. La red neuronal profunda, al considerar la coherencia temporal de la vibración logra aprender automáticamente patrones que le permiten clasificar las fallas y la localización de la mismas con una exactitud del 100 %. Cabe destacar el potencial de las

redes neuronales profundas que logran sobrepasar en exactitud los métodos convencionales de extracción de características.

Las investigaciones anteriormente expuestas permiten argumentar que la descomposición modal empírica entrega una muy útil herramienta para un proceso no estacionario y no lineal como el análisis de señales de vibraciones de rodamientos. Sumado a esto, la extracción de características de las funciones modales intrínsecas logran entregar una serie de valores representativos del fenómeno, los que se obtienen gracias a métricas relacionadas con la energía, entropía y forma de las funciones. Así, en la etapa de clasificación se utilizan múltiples tipos de clasificadores como ANN, DNN, SVM, ELM entre otras; donde una de las técnicas más novedosas y con mayores exactitudes son las relacionadas con algoritmos de Deep Learning.

3.2. Descomposición modal Empírica (EMD)

La técnica de Descomposición Modal Empírica (EMD) fue presentada por Huang et al. [14] en 1996, es considerada por los expertos una de las técnicas de análisis de señales tiempo-frecuencia más poderosas. El objetivo de la EMD es descomponer una señal en un conjunto de ondas mono modales y casi ortogonales, denominadas funciones modales intrínsecas (IMF), que representan las oscilaciones incrustadas en la señal. Las IMF son determinadas por la propia señal y no por una función kernel, esto permite a la EMD ser una técnica autoadaptable y adecuada para procesos no lineales y no estacionarios.

Para utilizar la EMD se debe considerar el supuesto que la señal $x(t)$ tiene al menos dos extremos o sólo puntos de inflexión que permiten encontrar por medio de la derivada sus extremos. De esta forma el proceso de descomposición comienza al identificar los máximos y mínimos de la onda, luego estos son conectados por medio de una interpolación cúbica, para crear la envolvente superior e inferior respectivamente. Con estas envolventes se debe realizar el cálculo del promedio, este se denominará m_1 , ahora la diferencia de m_1 con la señal $x(t)$ será la primera componente h_1 .

$$h_1 = x(t) - m_1 \tag{1}$$

De esta forma la componente encontrada h_1 , será una IMF C_1 si cumple dos condiciones. La primera condición es: En todo el conjunto de datos, el número de extremos y el número de interceptos con el eje cero debe ser de igual valor o diferir en uno. La segunda condición es: En cualquier punto del dominio de la señal, el valor del promedio definido por las envolventes superior e inferior debe ser igual a cero. En caso de no cumplir con las

condiciones se debe volver a iterar hasta tener éxito, pero la componente h_1 encontrada es considerada la nueva onda a descomponer.

Luego de finalizar el proceso de descomposición modal empírica de forma exitosa se obtiene desde la señal $x(t)$ un conjunto de M funciones modales intrínsecas C_i además de un residuo r_I .

$$x(t) = \sum_{i=1}^M (C_i + r_I) \quad (2)$$

Las funciones modales intrínsecas encontradas son oscilaciones incrustadas en la onda $x(t)$ que presentan diferentes bandas de frecuencias desde altas a bajas, mientras que el residuo r_I expresa la tendencia de la onda. La cantidad de funciones modales intrínsecas encontradas es determinada el propio algoritmo y depende de cada onda analizada[15].

Para garantizar que las IMF conservan suficiente sentido físico de las modulaciones tanto de amplitud como de frecuencia, debe existir un criterio para que el proceso de descomposición se detenga. Existen diversos criterios para la detención del algoritmo, donde su elección varía según la naturaleza del problema en análisis, por ejemplo una señal del mundo financiero tiene un comportamiento distinto que una vibración mecánica. Cabe destacar dos criterios que fueron propuestos por Huang et al.[14], el primero es el criterio de la desviación estándar. Este limita el valor de la desviación estándar, SD , el proceso es detenido cuando el valor sea menor que uno predeterminado, la fórmula se define como:

$$SD = \sum_{t=0}^T \frac{|h_{1(k-1)}(t) - h_{1k}(t)|^2}{h_{1(k-1)}^2(t)} \quad (3)$$

Esta fórmula obtiene la desviación estándar a partir de los dos resultados consecutivos de descomposición de la señal, donde h_{1k} y $h_{1(k-1)}$ representan la última y penúltima componente encontrada respectivamente.

El segundo es el criterio del residuo S , el algoritmo es detenido cuando el residuo r_I de la fórmula (2) se convierte en una función monótona de la cual no se puede extraer más funciones modales intrínsecas. Vale decir, si el residuo contiene más de un extremo se debe continuar la descomposición de la señal, en caso contrario el algoritmo debe detenerse.

Finalmente, el algoritmo de descomposición modal empírica propuesto en la investigación de Norden Huang[14] es descrito en la tabla 1.

La descomposición modal empírica tiene algunas debilidades, dentro de las más relevantes podemos mencionar dos. Primero, la elección de un criterio para detener el algoritmo, el cual influye en la cantidad de IMF que

Tabla 1: Algoritmo EMD

Algoritmo EMD

1. Inicializar: $r_0 = x(t); i = 1$
2. Extraer la i -ésima IMF C_i
 - a) Inicializar $h_{i(k-1)} = r_{i-1}; k = 1$
 - b) Extraer el máximo y mínimo local de $h_{i(k-1)}$
 - c) Interpolarse el máximo y mínimo local por "cubic spline" para formar la envolvente superior e inferior respectivamente de $h_{i(k-1)}$
 - d) Calcular la media $m_{i(k-1)}$ de las envolventes superior e inferior de $h_{i(k-1)}$
 - e) Dejar $h_{ik} = h_{i(k-1)} - m_{i(k-1)}$
 - f) Si h_{ik} es una IMF entonces establecer $C_i = h_{ik}$, sino volver al paso (b) con $k = k + 1$
3. Definir el resto $r_{i+1} = r_i - C_i$
4. Si r_{i+1} continua teniendo al menos 2 extremos entonces volver al paso 2 con $i = i + 1$ sino el proceso de descomposición a finalizado y r_{i+1} es el residuo de la señal.

intenta obtener desde la señal. Segundo, el problema de la mezcla de modos, que se define como una función modal intrínseca que contiene oscilaciones con escalas dramáticamente dispares. Esto se debe a la intermitencia en la onda y puede causar suavizamiento en la distribución tiempo frecuencia y podría dejar una IMF sin un significado físico, vale decir, que no describa una oscilación incrustada en la señal[15].

3.3. Descomposición modal Empírica Ensamblada (EEMD)

Uno de los inconvenientes de la Descomposición Modal Empírica es la aparición, en las Funciones Modales Intrínsecas, de señales con escalas ampliamente dispares o una señal de una escala similar que resida en diferentes componentes. Este problema se conoce como mezcla modal y es provocado por la intermitencia presente en la señal que se busca descomponer. La intermitencia en la señal tiene como efecto el aliasing en la distribución tiempo-frecuencia y que las IMF encontradas pierdan sentido físico, ya que no representan una oscilación incrustada sino una mezcla de estas.

En general, los datos capturados desde un fenómeno físico son una mezcla entre la señal correspondiente al fenómeno y un ruido originado por diversos factores. De esta forma, si una señal es capturada en diversas observaciones y cada una contiene un ruido diferente, el promedio de las observaciones representa de manera más exacta la señal en estudio. Con el objetivo de generalizar esta idea nace la Descomposición Modal Empírica Ensamblada [16], esta técnica utiliza los beneficios del ruido blanco para eliminar el problema de la mezcla modal.

La EEMD tiene como objetivo añadir diversos ruidos blancos de amplitud finita a una señal, proporcionando así una escala de referencia uniforme para facilitar el proceso de descomposición. De esta forma, a cada señal se añade un ruido blanco y luego es procesada por la EMD, este proceso es repetido N veces y al finalizar las IMF encontradas es el promedio de las N descomposiciones. Para reducir o eliminar el efecto de este ruido sobre las IMF se deben ensamblar N procesos de descomposición, produciendo así una cancelación entre los ruidos. En síntesis, este proceso logra a través del ruido blanco realizar una correcta separación de escalas de las oscilaciones presentes sin afectar las IMF encontradas por el algoritmo. Con base en la explicación previa, el algoritmo de Descomposición Modal Empírica Ensamblada presentada en la investigación de Zhaohua y Norden Huang [16] es sintetizado en la tabla 2.

Tabla 2: Algoritmo EEMD

Algoritmo EEMD

1. Inicializar el número de ensambladas M , la amplitud del ruido blanco añadido, y el número de ensayo $m = 1$
 2. Realizar el ensayo m -ésimo en la señal con ruido blanco añadido
 - a) Generar la serie de ruido blanco con la amplitud ya inicializada y añadir a la señal investigada $x_m(t) = x(t) + n_m(t)$ donde $n_m(t)$ indica el m -ésimo ruido blanco añadido y $x_m(t)$ representa la señal con ruido añadido del m -ésimo intento.
 - b) Descomponer la señal con ruido añadido $x_m(t)$ en I IMFs $C_{(i,m)}(i = 1, 2, \dots, I)$ utilizando el método EMD descrito previamente, donde $C_{(i,m)}$ denota la i -ésima IMF del m -ésimo intento, y I es el número de IMFs.
 - c) Si el número de intentos es más pequeño que el número requerido, $m < M$ entonces ir al paso (a) con $m = m + 1$. Repetir los pasos (a) y (b) nuevamente pero con la nueva serie de ruido blanco generada aleatoriamente en cada ocasión.
 3. Calcular la media de la ensamblada C_i de los M intentos por cada IMF $y_i = \frac{1}{M} \sum_{m=1}^M C_{(i,m)}$ $i = 1, 2, \dots, I, m = 1, 2, \dots, M$
 4. Informar la media $y_i(i = 1, 2, \dots, I)$ de cada una de los I IMFs como las IMFs finales.
-

A pesar que este nuevo método logra solucionar el problema de la mezcla modal mantiene los problemas de la EMD, la inexistencia de un criterio de detención adecuado y los efectos finales en las IMF encontradas[16]. Finalmente, tanto la cantidad de procesos de descomposición ensamblados como la amplitud del ruido añadido a la señal, son valores que deben ser ajustados en base a la señal analizada.

3.4. Deep Learning

Los algoritmos de Representation Learning permiten que una máquina sea alimentada con datos en bruto y automáticamente descubrir los patrones

necesarios para realizar una detección o clasificación. Las técnicas de Deep Learning son métodos de Representation Learning con múltiples niveles de representación, obtenidos de la unión de módulos simples pero no lineales. La idea principal es que cada nivel de representación, a través de funciones no lineales, transforme la data en bruto a un nivel más abstracto.

En labores de clasificación, principalmente de imágenes, los algoritmos de Deep Learning logran en las capas superiores de representación potenciar los aspectos relevantes para una correcta clasificación y eliminan las variaciones irrelevantes. Cabe señalar, que uno de los objetivos de estos algoritmos es reemplazar las características construidas manualmente por personas por eficientes algoritmos de aprendizaje no supervisados de características [17].

En el campo del diagnóstico de fallas, las señales de vibración etiquetadas son difíciles de adquirir y necesitan un ajuste experimental específico y elaborado. Por lo tanto, el aprendizaje de funciones y características sin supervisión puede proporcionar una solución eficaz para el diagnóstico de fallas. En este trabajo se utilizarán los Sparse Auto-Encoder uno de los típicos algoritmos de aprendizaje de características no supervisados.

3.4.1. Autoencoder

El autoencoder es un tipo de red neuronal de tres capas cuyo objetivo de salida es replicar los datos de entrada, esta forma de aprendizaje se conoce como aprendizaje no supervisado [18]. El autoencoder tiene dos funciones principales el codificador (encoder) y el decodificador (decoder), como se muestra en la figura 1 .

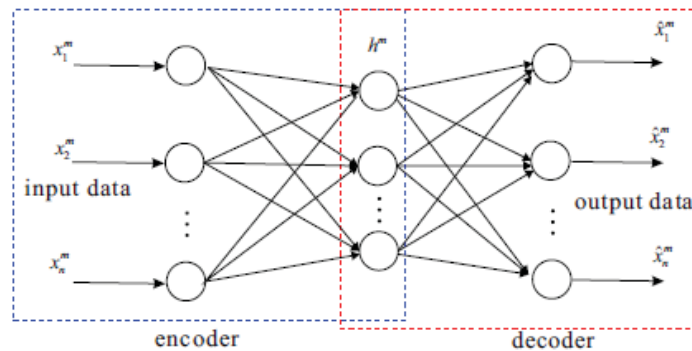


Figura 1: Arquitectura de un Auto Encoder[18]

La función del codificador transforma los datos de entrada a una representación comprimida, llevándolos desde un espacio de alta dimensión a un espacio de baja dimensión.

El proceso de codificación está definido por la siguiente función llamada $f_{encoder}$:

$$h^m = f_{encoder}(x^m) \quad (4)$$

Donde x^m es una muestra de los datos y h^m es el vector codificado obtenido de x^m .

La función del decodificador remodela la representación comprimida del paso anterior y reconstruye la entrada original. El proceso de decodificación está definido por la siguiente función llamada $f_{decoder}$:

$$\hat{x}^m = f_{decoder}(h^m) \quad (5)$$

Donde \hat{x}^m es el vector decodificado obtenido de h^m , el cual representa la reconstrucción de la entrada de datos.

Utilizar la red neuronal para replicar los datos de entrada se puede volver un proceso trivial, ya que esta puede simplemente aprender la función identidad. Ahora, con el fin de evitar este aprendizaje se puede aplicar una restricción a la capa oculta de la red, como limitar el número de neuronas. Esto debido a que los datos de entrada al atravesar la capa oculta, en el proceso de codificación, deberán ser representados con un menor número de neuronas. Así, se obtiene una representación comprimida de los datos de entrada que puede descomprimirse para reconstruir los mismos datos de entrada.

3.4.1.1. Autoencoder Sparse

Una restricción que es útil imponer a las unidades de la capa oculta es la de dispersión (Sparsity), esto con fin de obtener una representación comprimida que represente eficazmente la entrada. La restricción de sparsity desactiva neuronas en la capa oculta, así la red está forzada a generalizar y encontrar patrones en los datos que pueda representar mediante una cantidad menor de neuronas activas [19]. Bajo el supuesto que la función de activación de las neuronas es sigmoideal, una neurona está activa si su valor de salida es cercano a 1 o como inactiva si su valor de salida es cercano a 0. Ahora, a_j^2 denotara la activación de la unidad oculta j en el autoencoder. Sin embargo, esta notación no hace explícito cuál fue la entrada x que llevó a esa activación. Así, escribiremos $a_j^2(x)$ para denotar la activación de esta

unidad oculta cuando se le da a la red una entrada específica x . De esta forma:

$$\widehat{\rho}_j = \frac{1}{m} \sum_{i=1}^m a_j^2(x^i) \quad (6)$$

Sea $\widehat{\rho}_j$ la activación promedio de la unidad oculta j , promediada sobre el conjunto de datos de entrenamiento, el objetivo del autoencoder sparse es hacer cumplir la siguiente restricción:

$$\widehat{\rho}_j = \rho \quad (7)$$

Donde ρ es el parámetro de dispersión, típicamente un valor pequeño cercano a cero. En otras palabras, el objetivo es que la activación promedio de cada neurona j sea cercana a 0. Para cumplir este objetivo se debe añadir a la función del error de reconstrucción, un término que penalice la diferencia entre la activación promedio y la activación deseada de las neuronas. El término para lograr la penalización es la divergencia Kullback-Leibler entre una variable aleatoria Bernoulli con media ρ y una variable aleatoria Bernoulli con media $\widehat{\rho}_j$. La divergencia KL es una función estándar para medir cuán diferentes son dos distribuciones de probabilidad diferentes.

$$\sum_{i=1}^{s2} KL(\rho || \widehat{\rho}_j) \quad (8)$$

Donde:

$$KL(\rho || \widehat{\rho}_j) = \rho * \log \frac{\rho}{\widehat{\rho}_j} + (1 - \rho) \log \frac{1 - \rho}{1 - \widehat{\rho}_j} \quad (9)$$

Tal como se aprecia en la fórmula 9, la función de penalización tiene la propiedad de $KL(\rho || \widehat{\rho}_j) = 0$ si $\rho = \widehat{\rho}_j$ y en otro caso esta aumenta monótonamente como $\widehat{\rho}_j$ diverja de ρ . Por lo tanto la minimización de este término de penalización, tiene como efecto que el valor de $\widehat{\rho}_j$ esté cercano al valor de ρ .

De esta forma la función de costo del auto encoder sparse queda de la siguiente forma:

$$J_{sparse}(W, b) = J(W, b) + \beta KL(\rho || \widehat{\rho}_j) \quad (10)$$

El término $\widehat{\rho}_j$ implícitamente depende del peso oculto W y sesgo b del nodo oculto, porque es la activación promedio de la unidad oculta j , y la activación de una unidad depende de los parámetros (W, b) . Así $J(W, b)$ debe considerar

los valores de los pesos ocultos en la función del error de reconstrucción, de esta forma.

$$J(W, b) = \left[\frac{1}{N} \sum_{i=1}^N \left(\frac{1}{2} \|h_{W,b}(x^i) - y^i\|^2 \right) \right] + \frac{\lambda}{2} \sum_{l=1}^{m_l-1} \sum_{i=1}^{s_l} \sum_{j=1}^{s_{l+1}} (W_{ji}^l)^2 \quad (11)$$

Donde $J(W, b)$ es el error cuadrático medio sumado a la regularización de los pesos del autoencoder, donde λ es el coeficiente que afecta la regularización de los pesos. Así, β controla la importancia relativa del término de sparsity. Finalmente para encontrar los pesos ocultos que entreguen el menor error de reconstrucción, se debe minimizar la función de costo (11) utilizando algún algoritmo de optimización.

3.5. Clasificador Softmax

La función Softmax es una generalización de una función logística que transforma un vector z de valores reales de tamaño N a un vector $\sigma(z)$ de tamaño N de valores reales entre $[0, 1]$ que suman 1. Esta salida de la función 12 permite ser utilizada para representar la probabilidad que los valores del vector z pertenezcan a N posibles clases.[20]

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^N e^{z_k}} \quad (12)$$

El modelo de fallas propuesto requiere de un mecanismo para poder determinar frente a que clase de falla se encuentra, para esto se utiliza la función softmax. De esta forma, la utilización de una capa de la red neuronal profunda con neuronas que utilizan la función softmax permite una clasificación multiclase correcta.

4. Modelo de diagnóstico de fallas propuesto

Esta investigación propone un novedoso modelo de diagnóstico de fallas que utiliza la descomposición modal empírica y el aprendizaje profundo. Este modelo está enfocado a diagnosticar 10 clases de fallas según la severidad y ubicación en el rodamiento. El modelo tiene tres grandes etapas: el pre procesamiento de la señal de vibración, la extracción de características de la señal y el clasificador inteligente de la falla. A modo descriptivo el modelo se presenta como diagrama de bloques en la figura 2.

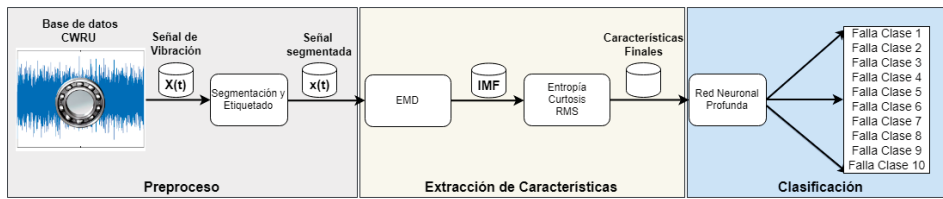


Figura 2: Diagrama de bloques modelo propuesto

4.1. Pre Proceso

El pre procesamiento de la señal consiste en cargar la señal de vibración al sistema y tomar M muestras de tamaño N , con esto se obtiene un conjunto de datos de M señales de menor tamaño. El tamaño de la muestra debe ser capaz de expresar el fenómeno en estudio, en este experimento se utiliza un tamaño de 800 puntos para cada una de las señales. En paralelo a cada una de estas muestras se asigna una clase binaria según el tipo y ubicación de falla.

4.2. Extracción de Características

La segunda etapa de extracción de características utiliza cada uno de estos segmentos y gracias a la descomposición modal empírica se obtienen las funciones modales intrínsecas incrustadas en la señal de vibración. Las oscilaciones encontradas por esta herramienta describen de forma más detallada el fenómeno en estudio, lo que permite obtener características locales de la falla presente en la señal. Luego de realizar la descomposición de la señal cada una de las oscilaciones es representada por medio de la curtosis, entropía y valor cuadrático medio.

4.2.1. Características utilizadas

Para representar cada una de las funciones modales intrínsecas encontradas por la descomposición modal empírica, se realiza el cálculo de tres características. Estas fueron seleccionadas gracias a su capacidad para capturar tanto fallas tempranas como de una alta severidad[21]. Los beneficios de utilizar estas características se pueden observar en análisis previos de fallas de elementos rotativos a través de los mismos [12],[22]. Las características calculadas a las IMF serán tres:

1. Curtosis: es una medida que sirve para analizar el grado de concentración que presentan los valores de una variable analizada alrededor de la zona central de la distribución de frecuencias. Su fórmula es:

$$curtosis = \sum_{i=1}^N \frac{(x_i - \bar{x})^4}{N\sigma^4} \quad (13)$$

2. Valor cuadrático medio: físicamente representa la potencia promedio de una señal, en esta aplicación se ha demostrado que tiene la habilidad de ser sensible a los impulsos de fallas [21]. El valor RMS de una señal $x(t)$ es calculado como la raíz cuadrada de la media aritmética de los cuadrados de los valores.

$$x_{rms} = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2} \quad (14)$$

3. Entropía: tiene relación con la incertidumbre que existe en cualquier experimento o señal aleatoria. Cuantifica el grado de irregularidad presente en una señal, si una señal es predecible o regular su entropía es mínima, en cambio si su entropía es máxima la señal es compleja de predecir e irregular.

$$H(X) = - \sum_{i=1}^n p(x_i) \log p(x_i) \quad (15)$$

4.3. Clasificador Inteligente

La tercera etapa utiliza las características de la etapa previa para alimentar el algoritmo de Deep Learning, que está compuesto de una red neuronal profunda de dos capas ocultas y una capa de clasificador softmax. Las capas ocultas de la red son pre entrenadas por medio de Autoencoders Sparse con el fin de encontrar los pesos ocultos. Luego de encontrados los pesos, se realiza el entrenamiento de la capa de clasificación, la que utiliza la función

Softmax (12), que se encargará de clasificar la localización y severidad de fallo. Para terminar el proceso de entrenamiento, se realiza el ajuste fino de la red neuronal profunda lo que permite ajustar los pesos y lograr un mejor diagnóstico de las fallas.

De forma general el algoritmo es como se describe en la tabla 3.

Tabla 3: Algoritmo Propuesto

Algoritmo Propuesto

1. Los datos de cada señal de vibración se divide en M segmentos pequeños de un determinado tamaño k .
 2. Estas señales de menor tamaño será procesadas por la descomposición modal empírica como se describió previamente, obteniendo N funciones modales intrínsecas.
 3. A cada una de estas IMFs nuevas les realizaremos una extracción de las características, estas serán la Curtosis, entropía y valor cuadrático medio (RMS).
 4. Gracias al paso previo tendremos una cantidad de características obtenidas de $N * 3$. Estas serán la entrada para realizar un pre entrenamiento de los Auto-Encoders Sparse apilados con el fin de encontrar el valor de los pesos ocultos.
 5. Finalmente con los pesos previamente encontrados se realiza el ajuste fino de la red neuronal profunda y procesan los datos para encontrar el tipo y severidad de fallo
-

4.4. Métricas utilizadas

En este trabajo las métricas a utilizar son la exactitud y el f-score para evaluar el desempeño del modelo de detección de falla frente a las diversas configuraciones. Las fórmulas de los criterios anteriores son calculadas como sigue [23]:

1. Exactitud: corresponde al total de datos bien clasificados, sobre el total de datos clasificados

$$Exactitud = \frac{tp + tn}{tp + tn + fp + fn} \quad (16)$$

2. Precisión: es el número de resultados positivos correctos dividido por el número de todos los resultados positivos.

$$Precision = \frac{tp}{tp + fp} \quad (17)$$

3. Recall: es el número de resultados positivos correctos dividido por el número de resultados positivos que deberían haber sido devueltos.

$$Recall = \frac{tp}{tp + fn} \quad (18)$$

4. F-score: Es una medida de la exactitud de una prueba. Puede interpretarse como un promedio ponderado de la precisión y el recall, donde un f-score alcanza su mejor valor en 1 y el peor en 0. Considera tanto la precisión como el recall de la prueba para calcular la puntuación, así el F-Score está dado por:

$$FScore = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (19)$$

Donde:

- Verdaderos Positivos (tp): número de casos clasificados con éxitos. Ej: Rodamiento con falla
- Verdaderos Negativos (tn): número de casos clasificados con rechazo correctos. Ej: Rodamiento sano
- Falsos Positivos (fp): número casos de clasificados de falsas alarmas. Ej: Rodamiento sano clasificado como con falla.
- Falsos Negativos (fn): número casos de clasificados con falta. Ej: Rodamiento con falla clasificado como sano.

5. Presentación de resultados

La presente sección expone los resultados en la creación, entrenamiento y pruebas del sistema de diagnóstico de fallas propuesto. La primera parte describe el experimento desde el cual se obtuvieron las señales de vibración. Luego, serán detalladas las fallas que el modelo diagnosticará y el conjunto de datos con el cual se pondrá a prueba el sistema. Posteriormente, se explicará el ajuste de los parámetros del modelo, la mejor configuración encontrada y los resultados obtenidos tanto para la extracción por medio de la EMD como de la EEMD. Para verificar los beneficios de la utilización de la configuración propuesta, se realiza un experimento comparativo frente a una red de 1 autoencoder. Finalmente, se realiza la discusión de los hallazgos.

5.1. Descripción del experimento utilizado

Los datos de vibraciones del rodamiento provienen desde el experimento realizado por el Case Western Reserve University Bearing Data Center [24], donde se recolectaron en un ambiente controlado vibraciones de rodamientos normales y con fallas.

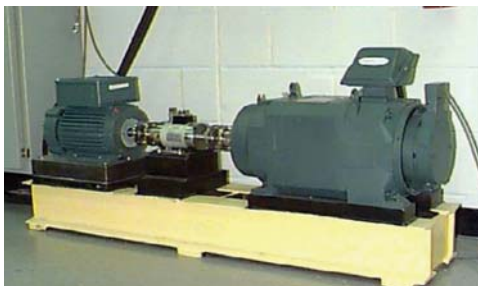


Figura 3: Foto de la configuración utilizada por el CWRU[24]

Como se muestra en la figura 3, el banco de pruebas está compuesto de un motor de 2 caballos de fuerza sostenido por rodamientos, un transductor de torque, un dinamómetro y electrónica de control. Las fallas de punto único se introdujeron en los rodamientos de prueba por medio de un mecanismo de electro-descarga con diámetros de falla de 7 mils, 14 mils, 21 mils, 28 mils y 40 mils (1 mil = 0.001 pulgadas). Los datos de vibración fueron capturados gracias a acelerómetros a 12.000 y 48.000 muestras por segundo, bajo 0,1,2 y 3 caballos de fuerza, la velocidad y potencia se registraron con un transductor de torque. [24]

5.2. Conjunto de datos utilizados

El conjunto de datos utilizado fue extraído desde el experimento previamente descrito, este conjunto cuenta con 40 archivos de vibración. Los archivos fueron organizados en 1 conjunto de datos para realizar el entrenamiento y prueba del sistema de diagnóstico creado. El conjunto de datos cuenta con todas las severidades de falla, velocidades y carga. La finalidad de esto es comprobar la eficacia del modelo frente a variaciones de velocidad y carga, con el fin de representar un ambiente real de funcionamiento de una máquina.

Tabla 4: Conjunto de datos utilizado

Identificador	Carga (HP)	Velocidad (RPM)	Cantidad de Clases
Dataset Mixed	0-1-2-3	1797-1772-1750-1730	10

Cada conjunto de datos tiene 10 clases de fallos según su ubicación y severidad, estas fueron etiquetadas binariamente en la etapa del pre procesamiento de las señales de vibración, como se muestra en la tabla 5.

Tabla 5: Clasificación de cada tipo de falla

Localización de la falla	Diámetro (mils)	Tipo de falla
Rodamiento sano	-	1
Falla en la bola	7	2
	14	3
	21	4
Falla en el anillo interno	7	5
	14	6
	21	7
Falla en el anillo externo	7	8
	14	9
	21	10

De esta forma, cada una de las muestras contiene una oscilación similar a las que se pueden apreciar en la figura 4. Donde se observa la particularidad de cada oscilación por medio de su amplitud en el tiempo, esto debido a que el fenómeno en estudio presenta una modulación en amplitud debido la carga a la cual está sometida el rodamiento. Sumado a esto las señales tienen estricta relación con la geometría y velocidad del rodamiento.

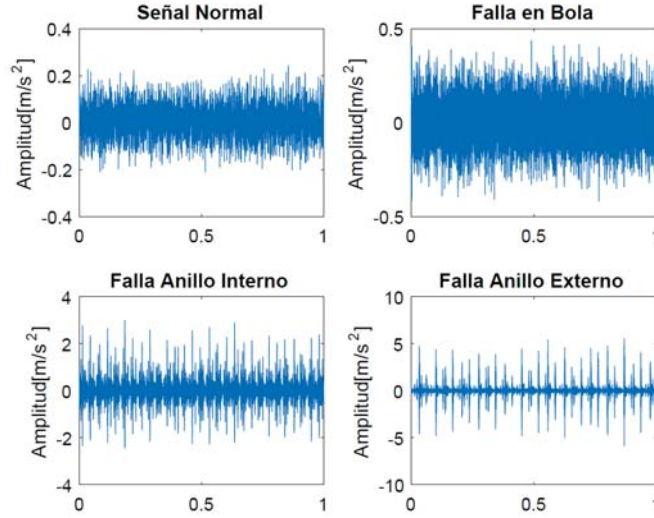


Figura 4: Señales de Vibración

5.3. Ajustes del modelo propuesto

5.3.1. Pre procesamiento

Antes de comenzar el proceso de entrenamiento de la red neuronal profunda es necesario formar la base de datos con las señales de vibración. Fueron probados diversos tamaños para realizar la segmentación de cada una de las señales, ya que el tamaño que se selecciona estará directamente relacionado con la cantidad de vueltas que realiza el rodamiento. Se debe considerar que en la mayor velocidad de rotación a 1797 RPM y por la frecuencia de muestreo de 12.000 muestras por segundo, una vuelta del rodamiento son aproximadamente 401 puntos de la señal.

Tabla 6: Exactitud en diversos tamaños de segmentación

Tamaño del segmento	Máximo	Mínimo	Promedio
400	96,22 %	94,67 %	95,4 %
800	98,11 %	96,56 %	97,23 %
1200	97,58 %	96,08 %	96,91 %

El tamaño óptimo para segmentar la señal fue encontrado en pruebas de múltiplos de 400 en diversas iteraciones, como se muestra en la tabla 6 el

tamaño de 800 puntos entrega la mejor exactitud promedio. Este tamaño en la menor velocidad de 1730 RPM representa 2 vueltas aproximadamente del rodamiento, mientras que en la mayor velocidad de 1797 RPM representa 1.9 vueltas del rodamiento. En este sentido y para tener igual cantidad de muestras en todas las señales de vibración, se tomaron 120.000 puntos de cada señal, así se obtienen 150 muestras de tamaño 800 puntos por cada señal. Junto a lo anterior, también se probaron diversos tamaños de muestras para entrenamiento y prueba del modelo, tal como se muestra en la tabla 7. La configuración con mayor exactitud promedio es utilizar un 70 % del conjunto de datos para entrenamiento y 30 % para pruebas.

Tabla 7: Tamaño del conjunto de datos y sus exactitudes

Porcentaje del dataset	No. Muestras Entrenamiento	No. Muestras Prueba	Exactitud Máxima (%)	Exactitud Mínima (%)	Exactitud Promedio (%)
50 %	3000	3000	97,57	96,33	96,4
60 %	3600	2400	97,96	96,42	97,11
70 %	4200	1800	98,11	96,56	97,23

5.3.2. Extracción de características

Luego de obtenida la base datos con las señales de vibración segmentadas, cada uno de estos segmentos es procesado por la descomposición modal empírica y la descomposición modal empírica ensamblada.

El programa utilizado para la EMD fue implementado por el investigador Patrick Flandrin, gracias a mostrar en la práctica un excelente desempeño y ser utilizado en diversas investigaciones [25]. De esta forma, se obtiene para señales sin fallas 6 funciones modales intrínsecas y, para señales con fallas, 8 o más funciones modales intrínsecas, tal como se observa en las figuras 5.

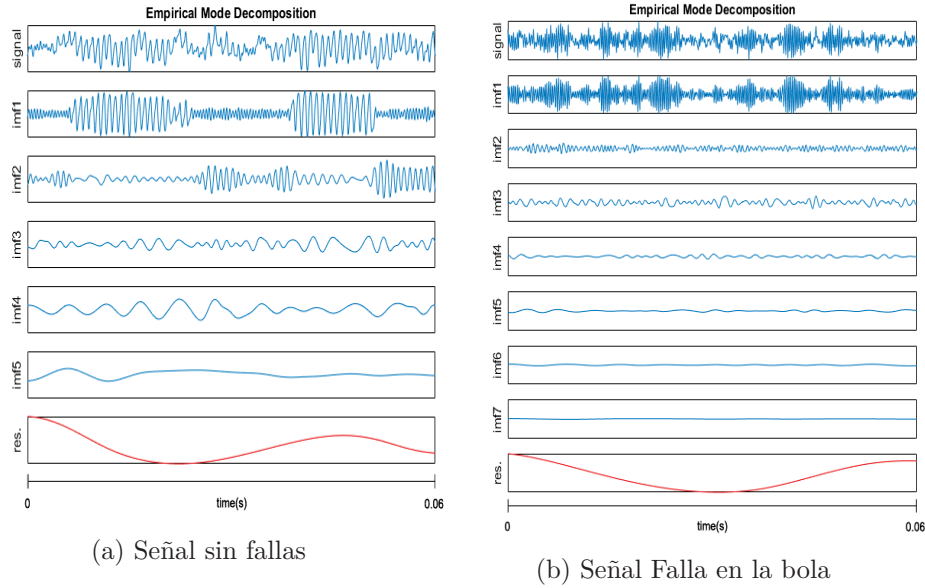


Figura 5: EMD diferentes señales

Para el procesamiento de las señales por medio de la descomposición modal empírica ensamblada, se utilizó el código de María Torres [26] el cual utiliza la EMD creada por Patrick Flandrin. El número de descomposiciones ensambladas utilizado fue de 30 debido a presentar los mejores resultados. Cabe señalar que se realizaron pruebas con 10, 50 y 100 ensambladas. En cuanto a la amplitud del ruido blanco añadido se probaron valores entre 0,2 y 0,8, donde las mejores exactitudes obtenidas fue con una amplitud de 0,6 como se puede observar en la tabla 8.

Tabla 8: Exactitud bajo distintos ruidos blanco

Noise Amplitud	Accuracy Max.(%)	Accuracy Min.(%)	Accuracy Prom.(%)
0,2	98,5	96,67	97,91
0,6	98,83	97,67	98,27
0,8	98,11	96,94	97,54

Con la mejor amplitud encontrada se procede a descomponer cada una de las muestras de la señal. Así, se obtuvieron para la señal saludable 9 IMF mientras que para la señal con fallas 10 o mas IMF. Estas oscilaciones encontradas se observan en las figuras 6. Este aumento en las funciones mo-

dales intrínsecas encontradas se explica quizás a una mezcla de oscilaciones presentes al descomponer por la EMD la señal de vibración.

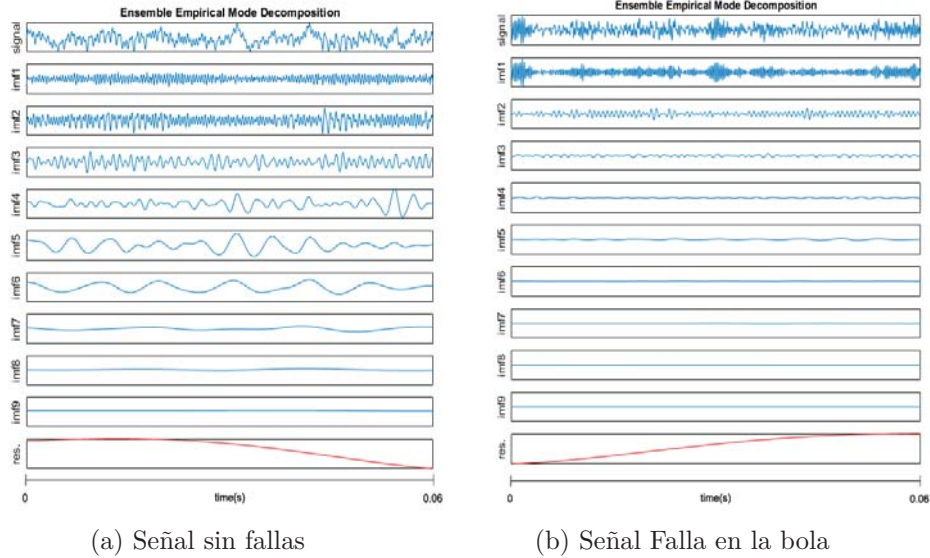


Figura 6: EEMD diferentes señales

La siguiente etapa consiste en representar a través de la curtosis, entropía y el rms, cada una de estas funciones modales intrínsecas, lo que permite obtener un nuevo conjunto de datos con el que se alimentará la DNN. En este sentido, luego del cálculo de las características, se obtiene un conjunto nuevo de datos de 6000 muestras por 18 y 27 características, ya que fue utilizada la menor cantidad de IMF. Cabe mencionar que para obtener un mejor rendimiento del algoritmo de aprendizaje, los valores obtenidos de las características debieron ser normalizados por el valor del mayor elemento de cada una de las características.

5.3.3. Aprendizaje profundo

El algoritmo de aprendizaje de Deep Learning para la clasificación de fallas tiene diversos parámetros. Estos requieren ser ajustados con base en los datos utilizados para obtener la mejor configuración de la DNN. Los parámetros relacionados con el autoencoder fueron ajustados por medio de valores aleatorios, distribuidos normalmente entre cierto intervalo. El primero de estos parámetros que se ajustó fue la cantidad de nodos ocultos de las capas ocultas, donde se encontró que la configuración con mejor exactitud

como se observa en la figura 7 es de 32 nodos para la primera capa y 16 nodos para la segunda.

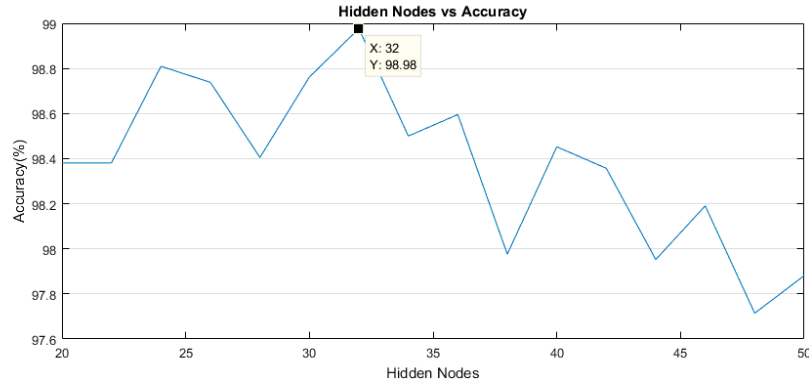


Figura 7: Exactitud vs Nodos ocultos

Sumado a lo anterior se probaron distintas relaciones en el número de nodos ocultos de los autoencoders. La relación entre el primer AE y del segundo AE fue 1 y 1, 1 y $\frac{1}{2}$, 1 y $\frac{1}{3}$, obteniendo la mejor exactitud con una relación de 1 y $\frac{1}{2}$. De esta forma, la configuración de la red es como se aprecia en la figura 8, 18 nodos de entrada, 32 nodos en la primera capa oculta y 16 en la segunda, un clasificador softmax y 10 nodos de salida.

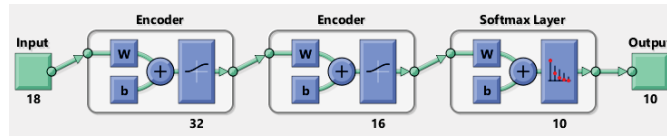


Figura 8: Red Neuronal Profunda utilizada

Las N capas ocultas de la red neuronal son pre entrenadas por medio de N autoencoders sparse, donde los pesos encontrados por el N-ésimo encoder son asignados a la N-ésima capa oculta de la red neuronal. La representación comprimida de las características, encontrada por el autoencoder 1 son las que alimentan el autoencoder 2 y las obtenidas de este último alimentan la capa de clasificación softmax.

Una de las variables que afecta el proceso de aprendizaje en la red es la proporción de sparsity ρ , explicada en la fórmula 7, la que afecta directamente qué tan especializada se puede volver una neurona. Donde para

encontrar el valor óptimo se realizaron pruebas con la configuración previamente encontrada y se varió el valor de la proporción desde 0.05 hasta 0.5. El valor óptimo fue de 0,4191, como se aprecia en la figura 9. Cabe señalar que se seleccionó este valor debido a repetirse en la realización de la misma prueba en diversos intentos.

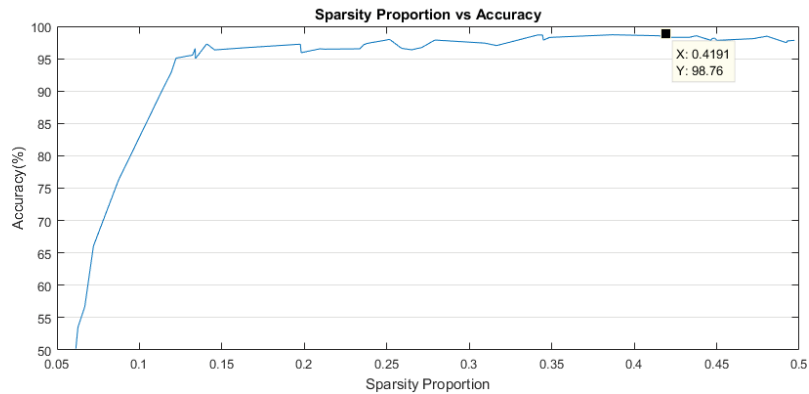


Figura 9: Proporción de Sparsity vs Exactitud

El valor del coeficiente β que multiplica la regularización de sparsity del autoencoder, fue ajustado entre 0.9 y 0.99 con base en los valores descritos en la investigación de Geoffrey Hinton [27]. De esta forma, se encontró que el valor que entrega la mejor exactitud es de 0.9033 tal como se aprecia en la figura 10.

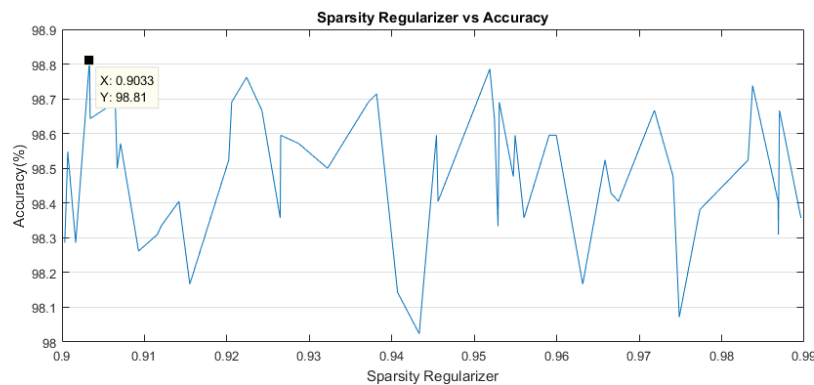


Figura 10: Coeficiente beta vs Exactitud

De igual forma que el parámetro β , el parámetro λ que es el coeficiente que multiplica la regularización de los pesos fue calculado con valores entre 0.0005 y 0.01. Cabe señalar que el valor de λ controla la importancia relativa de la penalización de los pesos, de esta forma es evitado el sobreentrenamiento. La figura 11 presenta que el coeficiente que entregó mejores resultados es 0.000757 para ambos autoencoders.

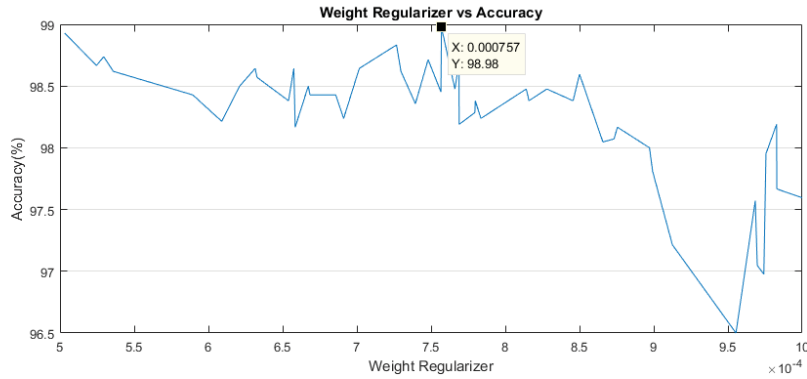


Figura 11: Coeficiente Lambda vs Exactitud

5.4. Primer Experimento

El presente experimento apunta a comparar la exactitud y eficiencia de la extracción de características a través de la EMD y la EEMD bajo el mismo algoritmo de Deep Learning previamente ajustado. Así, el sistema apunta a encontrar la existencia, localización y severidad de la falla bajo distintas velocidades y cargas. Con esto el sistema busca clasificar 10 tipos de clases tal como se describen en la tabla 5, por medio del conjunto de datos descrito previamente en la tabla 4.

Para realizar entrenamiento se utilizó el 70 % del total de muestras; para pruebas, el 30 % restante. La decisión se fundamenta en que el valor entregaba mejores exactitudes en repetidas ejecuciones, independiente de que se realizará con diferentes conjuntos de datos. Con el fin de obtener resultados consistentes, se realizaron 40 pruebas donde por cada prueba es reordenado el conjunto de datos aleatoriamente. Luego es entrenado y probado el sistema y almacenados sus resultados de exactitud, tiempos y f-score.

5.4.1. Extracción de características con EMD

El modelo, gracias a la extracción de características por medio de la EMD, alcanza una exactitud promedio de 97.23%, un máximo de 98.11% y un mínimo de 96.56%. Los valores de las pruebas realizadas se pueden observar en la figura 12.

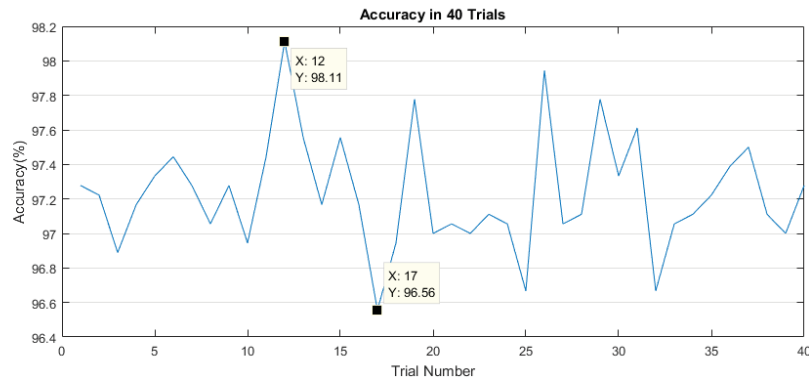


Figura 12: EMD - Exactitud en 40 intentos

En cuanto al tiempo, la extracción de características tarda en promedio 4.8 minutos en procesar las 6000 muestras, esto quiere decir que tarda aproximadamente 0.04 segundos en procesar cada segmento de la señal. El tiempo promedio en entrenamiento y testing del modelo es de 84.08 y 0.03 segundos respectivamente.

Con la finalidad de realizar un análisis con mayor detalle se calculó el f-score de cada una de las pruebas, para así poder observar cuáles son las clases que presentan mayor dificultad al momento de ser clasificadas. Se tomó el promedio por clase del f-score de las 40 pruebas realizadas, esto queda detallado en la figura 13.

La clase con menor f-score promedio es la falla de 0.021 mils en el anillo interno, seguida por la falla en el anillo interno de 0.014 mils. Este comportamiento indicaría que la falla no tiene una representación suficientemente estricta, por medio de las características calculadas. Como resultado de una baja representación, la red neuronal profunda no logra encontrar patrones para distinguir la falla de 0.021 mils en el anillo interno frente a otras clases.

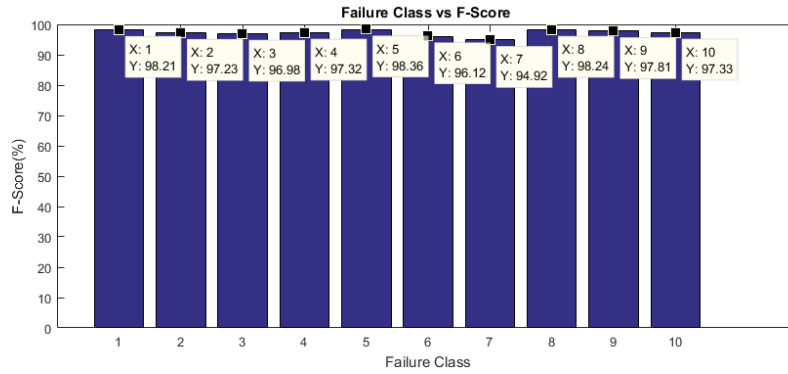


Figura 13: EMD - F-score promedio por clases en 40 intentos

5.4.2. Extracción de características con EEMD

Para analizar los beneficios de la EEMD en el procesamiento de señales, se realizaron las pruebas al modelo con la EEMD, donde alcanzó una exactitud promedio de 98.36 %, un máximo de 98.94 % y un mínimo de 97.56 %. Los valores de las pruebas realizadas se pueden observar en la figura 14.

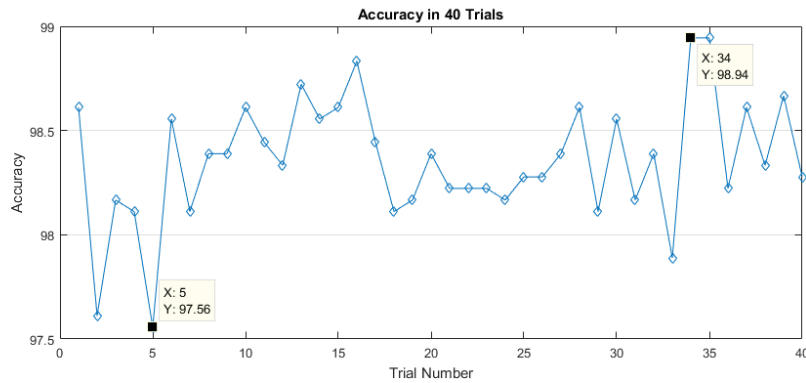


Figura 14: EEMD - Exactitud en 40 intentos

En cuanto al tiempo, la extracción de características tarda en promedio 137 minutos (2.2 horas) en procesar las 6000 muestras, esto quiere decir que tarda aproximadamente 1.36 segundos en procesar cada muestra de la señal. El tiempo en entrenamiento del modelo es en promedio 93.89 segundos, mientras que el tiempo promedio de testing es de 0.0372 segundos. Al igual

que con la extracción previa, se realizó el cálculo del f-score de cada una de las pruebas. El promedio de esto se puede ver en la siguiente figura 15.

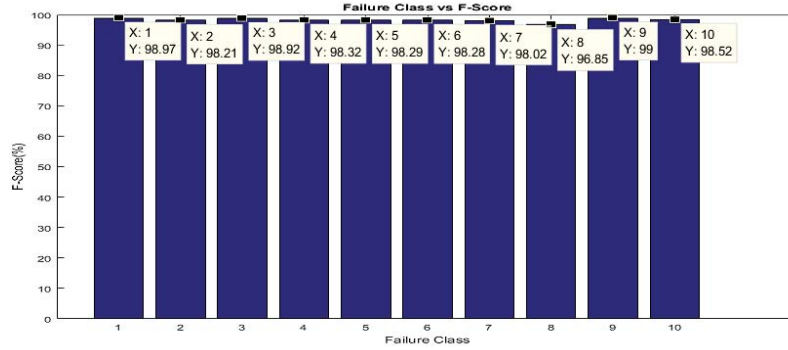


Figura 15: EEMD - F-score promedio por clases en 40 intentos

El F-Score más alto es la clase correspondiente a la falla del anillo externo de 0.014 mils, mientras que el menor F-Score es en la falla de anillo externo de 0.007 mils. Sumado a esto, no se aprecia una tendencia a disminuir la exactitud con respecto al aumento de la severidad de la falla.

5.5. Segundo Experimento

En este segundo experimento se pone a prueba el algoritmo de aprendizaje previamente utilizado frente a una red neuronal con solo una capa oculta pre entrenada con un autoencoder Sparse. Este experimento apunta a verificar la efectividad y eficiencia de utilizar un algoritmo basado en el aprendizaje profundo frente a una red neuronal pre entrenada por un autoencoder. Es importante destacar que el funcionamiento es igual al de la DNN utilizada en el primer experimento y solo fue variada la cantidad de capas ocultas, eliminando la característica de ser una red neuronal profunda.

El primer paso para realizar esta comparación fue ajustar los parámetros del autoencoder a utilizar, estos fueron ajustados de la misma forma que en el primer experimento. Los mejores parámetros encontrados fueron: 34 nodos ocultos, la proporción de sparsity $\rho = 0,4245$, el coeficiente $\beta = 0,9485$ y el coeficiente $\lambda = 0,0096$. Así la red neuronal cuenta con una estructura de 18 nodos de entrada, 34 nodos ocultos y 10 nodos de salida tal como se observa en la figura 16.

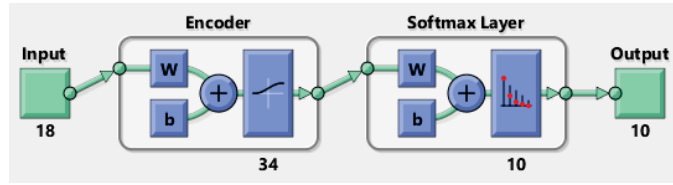


Figura 16: Red Neuronal 1 AE utilizada

De este modo se realizaron 40 pruebas donde se alcanzó una exactitud promedio de 98.19 %, una máxima del 98.94 % y una mínima del 97.06 %, tal como se observa en la figura 17. En cuanto al tiempo de entrenamiento, en promedio tarda 46.68 segundos y en el proceso de testing 0.0315 segundos, esto debido al menor costo computacional por la utilización de 1 autoencoder.

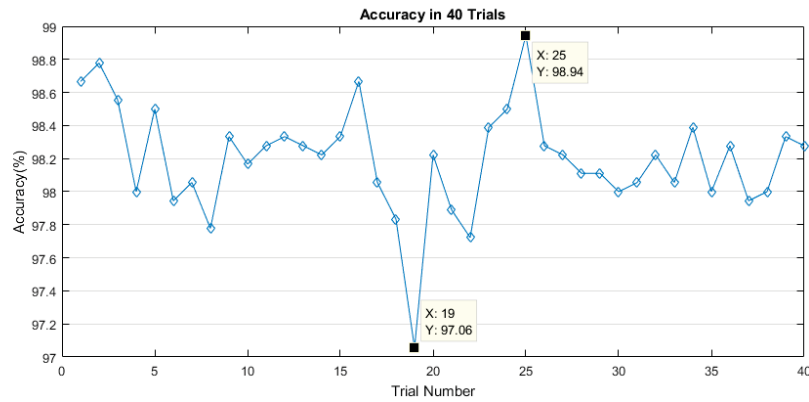


Figura 17: Exactitud en 40 intentos

Para observar el desempeño por cada clase de falla fue calculado el F-Score de cada una de las pruebas, en la figura 18 se muestra el promedio de los 40 resultados. El mejor F-Score promedio pertenece a la falla del anillo externo de 0.007 mils, mientras que el peor corresponde a la falla de 0.021 mils del anillo interno. Los resultados del F-score muestran un comportamiento similar que al utilizar 2 autoencoders, esto debido a la extracción de características utilizada y no el algoritmo de aprendizaje.

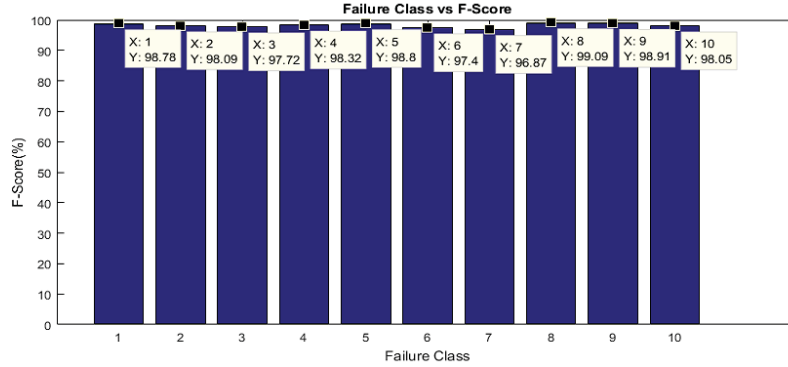


Figura 18: F-score promedio por clases en 40 intentos

Finalmente la tabla 9 resume los resultados obtenidos y su comparación tanto en exactitud como en tiempo de las diferentes configuraciones. Tal como se muestra la utilización de un autoencoder supera tanto en complejidad computacional y en exactitud a las redes neuronales profundas de dos y tres autoencoders.

Tabla 9: Comparación distintas configuraciones

		Exactitud (%)			T.Extracción(seg)			T.Train(seg)			T.Test(seg)			TOTAL(min)		
		Max.	Min.	Prom.	Max.	Min.	Prom.	Max.	Min.	Prom.	Max.	Min.	Prom.	Max.	Min.	Prom.
EMD	1 AE	98,94	97,06	98,19				63,58	46,53	48,68	0,05	0,03	0,03	6,37	5,39	5,70
	2 AE	98,11	96,56	97,23	319	277	293	91,17	81,31	84,08	0,04	0,03	0,03	6,83	5,97	6,29
	3 AE	96,83	94,94	95,88				159,79	95,99	113,20	0,09	0,03	0,04	7,98	6,22	6,78
EEMD	1 AE	99,56	98,44	99,09				81,07	62,49	71,65	0,04	0,02	0,03	143,58	129,79	138,19
	2 AE	98,94	97,56	98,36	8534	7725	8220	120,00	86,72	93,89	0,08	0,03	0,04	144,23	130,19	138,56
	3 AE	98,50	96,89	97,81				104,77	100,57	101,18	0,04	0,03	0,03	143,98	130,42	138,68

5.6. Discusión de resultados

Con base en los resultados expuestos, se discutirán ciertos puntos relevantes para el desarrollo de esta investigación.

Con respecto al primer experimento, luego de ajustados los parámetros y estableciendo una configuración óptima para el algoritmo de aprendizaje y clasificación, se realizaron pruebas con la utilización de EMD y EEMD. Con base en la exactitud, la EEMD logra mejores resultados que la EMD, aunque la diferencia no es superior a 1,1 puntos porcentuales aproximadamente. Este fenómeno se ve explicado en la mejor habilidad por parte de la EEMD para lograr descomponer la señal de vibración, tal como se demuestra en la literatura especializada [16]. No obstante, a pesar de superar a la EMD el

costo computacional de la EEMD es mayor debido a las ensambladas del algoritmo. De esta forma, la EEMD, a pesar de lograr mayor exactitud en un ambiente real, no es recomendable por el alto costo computacional que requiere, además de no generar una gran diferencia en las exactitudes finales.

Gracias a los resultados del f-score por clase de falla, el modelo presenta un mejor rendimiento al clasificar fallos en su estado inicial. Sumado a lo anterior, gracias a la fisonomía del rodamiento, la clasificación empeora según lo alejado que se encuentre la fuente de la vibración del acelerómetro que capta la señal. Este fenómeno se visualiza en el menor f-score del anillo interno. Así, al generarse el impacto en la zona del fallo, la vibración debe atravesar la película de lubricante, la bola y el anillo externo para llegar al acelerómetro. Debido a lo previamente mencionado se produce una distorsión en la señal que con la extracción de características por medio de la EMD las lleva a perder la capacidad de representar el fenómeno con exactitud.

Con respecto al segundo experimento, los resultados de la tabla 9 indican una superioridad tanto en exactitud como en complejidad computacional por parte de la utilización de un autoencoder. Esto queda plasmado en los resultados donde logra en promedio mejor exactitud que las configuraciones compuestas por dos y tres autoencoders. Sumado a esto se obtiene mayor exactitud máxima con un valor del 98.19%. En cuanto al F-Score obtiene mejores resultados que los otros algoritmos para todas las clases, pero siguiendo el mismo comportamiento previamente señalado debido a la extracción de características. Finalmente, se puede afirmar que la utilización de dos o más autoencoders, utilizando esta extracción de características, no logra entregar beneficios ni en exactitud ni complejidad, siendo la configuración de un autoencoder una mejor solución.

6. Conclusiones

Esta investigación presentó un modelo de diagnóstico de fallas de rodamientos basado en la descomposición modal empírica y en algoritmos de Deep Learning. El modelo apunta a detectar la existencia, severidad y localización de la falla bajo condiciones de velocidad y carga variable. Para verificar la exactitud del diagnóstico se utilizaron los datos recolectados por el CWRU, los cuales contienen señales de vibración de rodamientos bajo distintas condiciones de operación.

Con base en los resultados, el modelo propuesto EMD-DNN entrega una exactitud promedio de 97.23 % mientras que el modelo EEMD-DNN logra una exactitud promedio de 98.36 %. No obstante el costo computacional para procesar la señal de vibración por medio de la EMD es mucho menor que el requerido por la EEMD, el cual aumenta en razón de la cantidad de ensambladas utilizadas. De esta forma la EMD presenta mayores beneficios que la EEMD, a nivel de exactitud promedio versus el costo computacional.

En cuanto a la representación del fenómeno, la técnica de extracción de características utilizada logra representar de forma correcta los fallos en su etapa inicial. Sumado a esto el mejor F-Score promedio se observa en las clases de falla presentes en el anillo externo y bolas del rodamiento mientras que el peor F-Score se obtiene en el anillo interno. Independiente de la localización de la avería se observa que el diagnóstico empeora al aumentar la severidad de la falla.

Al comparar distintas configuraciones del clasificador, los resultados demostraron que una SLNN pre entrenada con un autoencoder sparse supera en exactitud y eficiencia computacional a una DNN. Sumado a esto, se requiere una gran cantidad de tiempo y experiencia para ajustar los parámetros de la red neuronal profunda, debido a tratarse de un problema multivariante que debe resolverse iterativamente. Con base en los resultados la utilización de un clasificador basado en una SLNN pre entrenada entrega un diagnóstico con mejor exactitud y a un menor costo computacional que una DNN.

La presente investigación con una exactitud promedio de 97.23 % logra superar a investigaciones similares como la de Wentao Mao[13]. Esta investigación alcanza un 76.67 % de exactitud utilizando la energía de las IMF junto a una SVM y un 82.8 % con la utilización de la descomposición por paquetes wavelet junto a una SVM. Esto demuestra el alto potencial descriptivo de la extracción de características utilizada y a una complejidad computacional muy baja frente a la descomposición de paquetes wavelet. No obstante, frente a técnicas más novedosas como la investigación de Ni-

baldo Rodriguez [8] que alcanza un 99.83 % de exactitud, no logra superar los resultados. Esto debido a la utilización de la extracción de características sofisticada por medio la transformada wavelet y la entropía de los valores singulares. Pese a lo anterior, cabe destacar el gran potencial del modelo propuesto para representar e identificar fallas en una etapa temprana y bajo condiciones de operación variable.

Queda en evidencia la gran capacidad de las descomposición modal empírica para descomponer señales no lineales y no estacionarias, lo que la convierte en una excelente solución al análisis de vibraciones. En este sentido, como objetivo futuro se pretende mejorar los resultados obtenidos por medio de la utilización de valores singulares aplicados a las funciones modales intrínsecas. Finalmente, el obtener resultados que superan investigaciones doctorales, da por cumplido el objetivo de crear un modelo adecuado para monitorear fallas en equipos críticos.

7. Referencias

- [1] SARABJEET SINGH, CARL Q. HOWARD, COLIN H. HANSEN, *An extensive review of vibration modelling of rolling element bearings with localised and extended defects*, Original Research Article, Journal of Sound and Vibration, Volume 357, 24 November 2015, Pages 300-330.
- [2] IEEE COMMITTEE REPORT, *Report of large motor reliability survey of industrial and commercial installation, Part I*, IEEE Transactions on Industry Applications, Volume 21, July 1985, Pages 853-864.
- [3] LOTFI SAIDI, JAOUHER BEN ALI, FARHAT FNAIECH, *Bi-spectrum based-EMD applied to the non-stationary vibration signals for bearing faults diagnosis*, Research Article, ISA Transactions, Volume 53, Issue 5, September 2014, Pages 1650-1660.
- [4] SAMANTA BISWANATH, AL-BALUSHI KHAMIS RAJAB *Artificial Neural Network Based Fault Diagnostics of Rolling Element Bearings Using Time-Domain Features*, Mechanical Systems and Signal Processing, Volume 17, Issue 2, March 2003, Pages 317-328
- [5] FENG JIA, YAGUO LEI, JING LIN, XIN ZHOU, NA LU *Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data*, Mechanical Systems and Signal Processing, Volumes 72–73, May 2016, Pages 303-315
- [6] YANG YU, YUDEJIE, CHENG JUNSHENG *A roller bearing fault diagnosis method based on EMD energy entropy and ANN*, Journal of Sound and Vibration, Volume 294, December 2005, Pages 269–277
- [7] XIAOYUAN ZHANG, JIANZHONG ZHOU *Multi-fault diagnosis for rolling element bearings based on ensemble empirical mode decomposition and optimized support vector machines*, Mechanical Systems and Signal Processing, Volume 41, Issue 1-2, December 2013, Pages 127-140
- [8] NIBALDO RODRIGUEZ, GUILLERMO CABRERA, CAROLINA LAGOS AND ENRIQUE CABRERA *Stationary Wavelet Singular Entropy and Kernel Extreme Learning for Bearing Multi-Fault Diagnosis*, Entropy 2017, Volume 19, Issue 10, October 2017, Pages 1-15
- [9] JAOUHER BEN ALI, NADER FNAIECH, LOTFI SAIDI, BRIGITTE CHEBEL-MORELLO, FARHAT FNAIECH *Application of empirical mode decomposition and artificial neural network for automatic bearing fault*

- diagnosis based on vibration signals*, Original Research Article, Applied Acoustics, Volume 89, March 2015, Pages 16-27.
- [10] YU YANG, DEJIE YU, JUNSHENG CHENG *A fault diagnosis approach for roller bearing based on IMF envelope spectrum and SVM*, Measurement, Volume 40, October 2006, Pages 943–950
- [11] RAN ZHANG, ZHEN PENG, LIFENG WU, BEIBEI YAO AND YONG GUAN *Fault Diagnosis from Raw Sensor Data Using Deep Neural Networks Considering Temporal Coherence*, Research Article, Sensors 2017, Volume 17 Issue 3,549, 9 March 2017, Pages 1-17.
- [12] JACEK DYBALA, RADOSLAW ZIMROZ *Rolling bearing diagnosing method based on Empirical Mode Decomposition of machine vibration signal*, Original Research Article, Applied Acoustics, Volume 77, 2014, Pages 195-203.
- [13] WENTAO MAO, JIANLIANG HE, YUAN LI Y YUNJU YAN *Bearing fault diagnosis with auto-encoder extreme learning machine: A comparative study*, Journal of Mechanical Engineering Science, 3 November 2016, Pages 1-19.
- [14] NORDEN E. HUANG, ZHENG SHEN, STEVEN R. LONG, MANLI C. WU, HSING H. SHIH, QUANAN ZHENG, NAI-CHYUAN YEN, CHI CHAO TUNG, HENRY H. LIU *The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis*, Proceedings of the royal society a mathematical, physical and engineering sciences, Volume 454, issue 1971, 8 March 1998, Pages 903-995.
- [15] YAGUO LEI., JING LIN., ZHENGJIA HE., MING J. ZUO. *A review on empirical mode decomposition in fault diagnosis of rotating machinery*, Mechanical Systems and Signal Processing, Volume 35, Issues 1–2, February 2013, Pages 108-126.
- [16] ZHAOHUA WU, NORDEN E. HUANG *Ensemble empirical mode decomposition: a noise-assisted data analysis method*, Advances in Adaptive Data Analysis, Volume 01, Issue 01, Enero 2009, Pages 1-49.
- [17] YANN LECUN, YOSHUA BENGI, GEOFFREY HINTON *Deep Learning*, Nature, Volume 521, 28 May 2015, Pages 436–444.
- [18] TAO MA, FEN WANG, JIANJUN CHENG, YANG YU AND XIAOYUN CHEN *A Hybrid Spectral Clustering and Deep Neural Network Ensemble*

- Algorithm for Intrusion Detection in Sensor Networks*, Sensors 2016, Volume 16, 13 Octubre 2016, Pages 1-23.
- [19] WENJUN SUN, SIYU SHAO, RUI ZHAO, RUQIANG YAN, XINGWU ZHANG, XUEFENG CHEN *A Sparse Auto-encoder-Based Deep Neural Network Approach for Induction Motor Faults Classification*, Measurement (2016), Volume 89, Julio 2016, Pages 171-178.
- [20] CHRISTOPHER M. BISHOP *Pattern Recognition and Machine Learning.*, Springer, New York, 2006 Pages 113-116.
- [21] ZHANGUO XIA, SHIXIONG XIA, LING WAN AND SHIYU CAI *Spectral Regression Based Fault Feature Extraction for Bearing Accelerometer Sensor Signals*, Sensors 2012, Volume 12, 12 Octubre 2012, Pages 13694-13719.
- [22] WEIHUA LI TIELIN SHI GUANGLAN LIAO SHUZI YANG *Feature extraction and classification of gear faults using principal component analysis*, Journal of Quality in Maintenance Engineering, Volume 9, Issue 2, 2003, Pages 132 – 143.
- [23] WIKIPEDIA *Precision and recall*, https://en.wikipedia.org/wiki/Precision_and_recall, Acceso 28/04/2017
- [24] CASE WESTERN RESERVE UNIVERSITY BEARING DATA CENTER *Data File*, <http://csegroups.case.edu/bearingdatacenter/pages/download-data-file>, Acceso 28/04/2017
- [25] GABRIEL RILLING, PATRICK FLANDRIN, PAULO GONÇALVES. *On empirical mode decomposition and its algorithms*, Proceedings of IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing NSIP-03, 2003, Grado, Italy.
- [26] MARÍA E. TORRES, MARCELO A. COLOMINAS, GASTÓN SCHLOTTHAUER, PATRICK FLANDRIN *A complete ensemble empirical mode decomposition with adaptive noise*, 36th IEEE International Conference on Acoustics, Speech, and Signal Processing, Article number 5947265, May 2011, Pages 4144-4147
- [27] GEOFFREY HINTON *A practical guide to training restricted boltzmann machines*, Lecture Notes in Computer Science, Lecture Number 7700, 2012, Pages 599-619.