



CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS
AVANZADOS DEL INSTITUTO POLITÉCNICO
NACIONAL

UNIDAD ZACATENCO
DEPARTAMENTO DE CONTROL AUTOMÁTICO

Diagnóstico de daño en estructuras civiles

Tesis que presenta

Ing. Odette Ortigoza Sánchez

Para obtener el grado de

Maestra en Ciencias

En la especialidad de

Control Automático

Directores de la Tesis:

Dr. Wen Yu

Dr. Jesús Morales Valdez

Ciudad de México

Febrero, 2024

Agradecimientos

Agradezco al CONAHCYT (Consejo Nacional de Humanidades Ciencias y Tecnologías) por la beca otorgada durante mis estudios de maestría.

A mi director y codirector de tesis, el Dr. Wen Yu Liu y el Dr. Jesús Morales Valdez, por la atención, tiempo brindado y constancia durante este proyecto.

También, hago mención de los sinodales que evaluaron esta tesis, cuyos comentarios me ayudaron a enriquecer mi trabajo.

Así como, a los investigadores del Departamento de Control Automático, ya que sus palabras fueron sabias, sus conocimientos rigurosos y precisos, gracias por su paciencia, por compartir sus conocimientos de manera profesional e invaluable, por su dedicación y perseverancia durante mis estudios.

Finalmente, a mi padre Manuel Ortigoza Campos, que me enseñó que aun en nuestros peores momentos podemos tener una razón para seguir adelante, y que con: determinación, paciencia y pasión a nuestra carrera todo es posible.

*Dedicado a mi madre
María de Lourdes Sánchez Cruz*

*Los que son amados no pueden morir,
porque amor significa inmortalidad.*

Emily Dickinson

Índice general

Agradecimientos	II
Resumen	VI
Abstract	VIII
Índice de figuras	X
Índice de tablas	XI
Nomenclatura	XIII
1. Introducción	1
1.1. Planteamiento del problema	2
1.2. Objetivo general	3
1.3. Justificación	3
1.4. Estructura de la tesis	4
1.5. Aportaciones	5
2. Estado del arte	6
2.1. ¿Qué es el diagnóstico estructural?	6
2.2. Tipos de diagnósticos estructurales	7
2.2.1. Rehabilitación de edificios	7
2.2.2. Adquisición de datos	7
2.3. Métodos de identificación	7
2.4. Métodos basados en frecuencia de vibración	8
2.4.1. Trabajos hasta el momento	9
2.4.2. Ventajas, limitaciones y desventajas	10
2.5. Métodos basados en análisis modal	10
2.5.1. Trabajos hasta el momento	11
2.5.2. Ventajas, limitaciones y desventajas	12
2.6. Métodos basados en propagación de onda	12
2.6.1. Trabajos hasta el momento	13
2.6.2. Ventajas, limitaciones y desventajas	13
2.7. Método basado en redes neuronales	14
2.7.1. Trabajos hasta el momento	14
2.8. Redes neuronales	16
2.9. Clasificación de Redes Neuronales	17
2.9.1. Clasificación según la topología de red	17

2.9.2. Clasificación según el método de aprendizaje	24
2.10. Comparación de redes neuronales	26
2.11. Redes neuronales y sus aplicaciones	29
2.11.1. Ventajas	30
2.11.2. Desventajas	30
3. Modelo matemático de las estructuras de tipo edificio	32
3.1. Conceptos fundamentales	32
3.1.1. Vibración	32
3.1.2. Movimiento armónico	33
3.1.3. Frecuencia	33
3.1.4. Amortiguamiento	34
3.2. Instrumentos de medición de vibraciones	34
3.2.1. Medición de desplazamiento	35
3.2.2. Medición de aceleración	36
3.2.3. Información requerida para la medición	36
3.3. Modelo matemático	37
3.3.1. Características que pueden ser modificadas	42
3.4. Modelo de histéresis de Bouc-Wen	44
4. Modelo implementado con Red Neuronal Multicapa	47
4.1. Estructura de una neurona artificial	47
4.1.1. Función de entrada	48
4.1.2. Función de activación	49
4.1.3. Función de salida	54
4.2. Estructura de una red neuronal multicapa	55
4.2.1. Capa de entrada	56
4.2.2. Capas ocultas	56
4.2.3. Capa de salida	56
4.2.4. Formas de conexión entre neuronas	57
4.2.5. Hiperparámetros	57
4.3. Red Neuronal Multicapa	57
4.3.1. Aplicaciones de redes neuronales multicapa	58
4.3.2. Aprendizaje por algoritmo de retropropagación (Backpropagation)	59
4.4. Evaluación del modelo implementado redes neuronales	64
4.4.1. Prototipo de estructura tipo edificio	64
4.4.2. Diseño del experimento	66
4.4.3. Diseño de red neuronal	69
4.4.4. Fase de aprendizaje para redes neuronales	71
4.5. Diagnóstico de daño con redes neuronales	75
4.5.1. Red neuronal monocapa	75
4.5.2. Evaluación con red neuronal multicapa	79
5. Conclusiones y trabajos futuros	83
5.1. Conclusiones	83
5.2. Trabajo futuro	84
A. Apéndice A: Características del prototipo tipo edificio	85

Resumen

En la presente tesis se desarrollan diferentes esquemas de diagnóstico con la finalidad de obtener un oportuno diagnóstico de daño en estructuras tipo edificios, que permita el monitoreo de la salud estructural para así asegurar la integridad de las mismas ante la presencia de movimientos sísmicos.

Entre las estrategias empleadas están los enfoques de métodos metaneurísticos, los cuales consisten en aproximaciones diseñadas para resolver problemas de optimización combinatoria. Este enfoque se basa en la capacidad de las redes neuronales para aprender patrones complejos a partir de datos estructurales, permitiendo una evaluación precisa y eficiente de posibles deterioros.

Se sabe que el territorio mexicano tiene una gran variedad de fenómenos meteorológicos, los cuales se encuentran presentes dentro de la atmósfera, estos pueden ser de un efecto positivo o negativo en el ecosistema, por lo que estos fenómenos naturales perjudican la salud estructural de los edificios; en especial los causados por el Cinturón de Fuego que es la zona sísmica donde se encuentra el país.

El diagnóstico de daño en estructuras civiles implica la identificación y evaluación de posibles defectos, deformaciones o desgastes que puedan comprometer la integridad estructural; en este contexto, las redes neuronales ofrecen una alternativa no invasiva, aprovechando la capacidad de procesamiento de información, lo cual es importante para los sistemas que procesan grandes cantidades de datos (por ejemplo, sistemas de radar y sonar), permitiendo emplearlas para diagnóstico en una estructura civil.

Los métodos físicos de identificación y evaluación de daño mencionados anteriormente, requieren datos previos; un ejemplo es la rehabilitación de edificios, el cual requiere diferentes análisis para saber el estado de la estructura, como lo son: análisis cualitativo y cuantitativo debido a que existen edificios con la misma edad, materiales y características similares al edificio que se estudie, todos estos análisis permiten recopilar la información necesaria para conformar una base de datos que permita tener la información del estado estructural de un edificio en específico.

Por lo que utilizando datos recopilados a través de tecnologías de monitoreo como: sensores de vibración, acelerómetros y otros dispositivos de detección. La metodología propuesta se basa en el aprendizaje de la red neuronal con un conjunto de datos representativos de condiciones estructurales normales y dañadas.

De modo que los métodos de identificación de daño, implican la observación y el análisis del un componente de un sistema a lo largo del tiempo, por lo cual se mencionaran diferentes métodos que se han utilizado, denotando sus ventajas y desventajas, así como el modo en el cual es empleado en las estructuras; y el porqué se combina con las redes neuronales.

Para ello, el modelo matemático implementado para el trabajo en cuestión, se tomaron los conceptos de las vibraciones, los movimientos armónicos, frecuencia y el amortiguamiento; con respectivas características de cada uno de ellos y con enfoques respectivos sobre las necesidades que se deben tener para la toma de datos que interpretarán y usarán las redes

neuronales.

El diagnóstico de daño estructural requiere el uso de técnicas avanzadas. Como lo son, las redes neuronales, las cuales se han convertido en una herramienta imprescindible para ingenieros y profesionales en el área.

De modo que se emplearan las redes neuronales multicapa, las cuales se encuentra compuesta por múltiples capas de neuronas, que procesaran los datos obtenidos de diferentes instrumentos de medición para la estructura a estudiar. Para así obtener un diagnóstico de daño estructural que se base en la recopilación y procesamiento de datos de diferentes escenarios de daños que puede presentar una estructura.

Por lo cual, el presente trabajo se embarca en una exploración integral de las aplicaciones de las redes neuronales, destacando su potencial en el diagnóstico de daños en estructuras civiles. De modo que este enfoque interdisciplinario entre la ingeniería estructural y la inteligencia artificial, aspira a propiciar avances que redefinan la forma en que abordamos la salud estructural de las infraestructuras.

Donde la combinación de tecnologías de monitoreo avanzadas con la capacidad de aprendizaje de máquinas ofrece una perspectiva innovadora para el futuro de la evaluación estructural, promoviendo la seguridad, eficiencia y sostenibilidad a largo plazo de nuestras infraestructuras civiles impulsando así un futuro más seguro y sostenible.

Abstract

In this thesis, different diagnostic schemes are developed with the purpose of obtaining a timely diagnosis of damage in building-type structures, which allows the monitoring of structural health to ensure their integrity in the presence of seismic movements.

Among the strategies used are metaheuristic method approaches, which consist of approaches designed to solve combinatorial optimization problems. This approach is based on the ability of neural networks to learn complex patterns from structural data, allowing accurate and efficient evaluation of possible deteriorations.

It is known that the Mexican territory has a wide variety of meteorological phenomena, which are present within the atmosphere, these can have a positive or negative effect on the ecosystem, so these natural phenomena harm the structural health of buildings; especially those caused by the Ring of Fire, which is the seismic zone where the country is located.

The diagnosis of damage in civil structures involves the identification and evaluation of possible defects, deformations or wear that may compromise structural integrity; In this context, neural networks offer a non-invasive alternative, taking advantage of the information processing capacity, which is important for systems that process large amounts of data (for example, radar and sonar systems), allowing them to be used for diagnosis in a civil structure.

The physical methods of identification and damage assessment mentioned above require prior data; An example is the rehabilitation of buildings, which requires different analyzes to know the state of the structure, such as: qualitative and quantitative analysis because there are buildings with the same age, materials and similar characteristics to the building being studied, all These analyzes allow us to collect the necessary information to form a database that allows us to have information on the structural state of a specific building.

So using data collected through monitoring technologies such as: vibration sensors, accelerometers and other detection devices. The proposed methodology is based on learning the neural network with a set of data representative of normal and damaged structural conditions.

So the damage identification methods involve the observation and analysis of a component of a system over time, for which different methods that have been used will be mentioned, denoting their advantages and disadvantages, as well as the way in which it is used in structures; and why it is combined with neural networks.

For this, the mathematical model implemented for the work in question, the concepts of vibrations, harmonic movements, frequency and damping were taken; with respective characteristics of each of them and with respective approaches on the needs that must be had for collecting data that will be interpreted and used by neural networks.

Diagnosing structural damage requires the use of advanced techniques. As are neural networks, which have become an essential tool for engineers and professionals in the area.

So that multilayer neural networks will be used, which are composed of multiple layers of neurons, which will process the data obtained from different measuring instruments for the

structure to be studied. In order to obtain a diagnosis of structural damage that is based on the collection and processing of data from different damage scenarios that a structure may present.

Therefore, this work embarks on a comprehensive exploration of the applications of neural networks, highlighting their potential in the diagnosis of damage to civil structures. So this interdisciplinary approach between structural engineering and artificial intelligence aspires to foster advances that redefine the way we approach the structural health of infrastructure. Where the combination of advanced monitoring technologies with machine learning capabilities offers an innovative perspective for the future of structural assessment, promoting the safety, efficiency and long-term sustainability of our civil infrastructures thus driving a safer and more sustainable future.

Índice de figuras

1.1. Monitoreo de salud estructural.	2
1.2. Placas tectónicas que afectan a la República Mexicana.	4
2.1. Red neuronal biológica.	17
2.2. Estructura red neuronal monocapa.	17
2.3. Estructura red neuronal convolucional.	19
2.4. Recorrido del filtro de acuerdo al número de pasos definidos.	20
2.5. Estructura red neuronal recurrente.	22
2.6. Estructura red neuronal LSTM.	24
3.1. Movimiento armónico.	33
3.2. Proceso de Medición de vibraciones.	35
3.3. Edificio de n - pisos.	39
3.4. Aumento de amortiguamiento por medio de masa sintonizada, en Torre Taipei 101.	43
3.5. Funcionamiento de un amortiguador viscoso.	44
4.1. Estructura de una neurona artificial.	48
4.2. Función sigmoide.	51
4.3. Función tangente hiperbólica.	52
4.4. Función ReLU.	53
4.5. Función Leaky ReLU.	53
4.6. Función Softmax.	54
4.7. Red neuronal multicapa totalmente interconectada.	56
4.8. Identificación de sistema 4.10, con red neuronal monocapa.	58
4.9. Identificación de sistema 4.10, con red neuronal multicapa.	59
4.10. Prototipo de edificio de 2 niveles localizado en el Laboratorio DCA.	64
4.11. Comportamiento de las componentes del registro sísmico de Northridge 1994.	65
4.12. Adquisición de datos empleando MATLAB 2012b/Simulink.	66
4.13. Energías con presencia de daño en el piso 2.	67
4.14. Deformación de la estructura.	68
4.15. Trayectoria en la que se realiza la conexión entre neuronas de tipo propagación hacia delante.	74
4.16. Trayectoria en la que se realiza la actualización de pesos sinópticos.	75
4.17. Resultados obtenidos al utilizar red monocapa con función de activación sigmoide.	77
4.18. Resultados obtenidos al utilizar red monocapa con daño en el piso 2.	78
4.19. Resultados obtenidos al utilizar red monocapa con daño en el piso 1 y piso 2.	78
4.20. Resultados obtenidos al utilizar red neuronal multicapa con daño en el piso 2.	81

4.21. Resultados obtenidos al utilizar red neuronal multicapa con daño en el piso 1 y piso 2. 81

Índice de tablas

2.1. Comparación entre una neurona biológica y una artificial.	17
2.2. Comparación entre diferentes tipos de redes neuronales.	28
3.1. Clasificación de modelos Bouc-Wen BIBO estables	45
4.1. Funciones de activación.	54
4.2. Experimentos propuestos para el prototipo tipo edificio.	68
4.3. Parámetros finales de la red neuronal monocapa	70
4.4. Parámetros finales de la red neuronal multicapa.	70
4.5. Ajuste de parámetros red monocapa.	76
4.6. Resultados obtenidos al utilizar red monocapa.	77
4.7. Ajuste de parámetros red neuronal multicapa.	79
4.8. Resultados obtenidos al utilizar red neuronal multicapa.	80
A.1. Características del prototipo tipo edificio de 2 pisos.	85
A.2. Escalamiento de movimiento sísmico Northridge para el prototipo tipo estructura.	86

Nomenclatura

SHM	Structural Health Monitoring (Monitoreo de Salud Estructural)
FM	Flexibility Method (Método de Flexibilidad)
SM	Stiffness Method (Método de Rigidez)
MA	Mixed Approach (Enfoque Mixto)
CWT	Continous Wavelet Transform (Transformada Continua Wavele)
FFT	Fast Fourier Transform (Transformada Rápida de Fourier)
NDE	Non-Destructive Evaluation (Evaluación no Destructiva)
SPC	Statistical Processes Control (Control Estadístico de Procesos)
DP	Damage Forecast (Pronóstico de Daños)
ISE	Interacción Suelo Estructura
ARMA	AutoRegressive Moving Average models (Modelos Autorregresivos de Media Móvil)
MLP	Multilayer Perceptron (Perceptrón Multicapa)
CNNFI	Convolutional Neural Networks for Indentification (Redes Neuronales Convolucionales para Identificación)
CNNFC	Convolutional Neural Networks for Classification (Redes Neuronales Convolucionales para Clasificación)
NN	Neural Network (Red Neuronal)
CNN	Convolutional Neural Networks (Red Neuronal Convolutacional)
RBG	Red, Green, Blue (Rojo, Verde, Azul)
RNN	Recurrent Neural Network (Red Neuronal Recurrente)
SRN	Simple Recurrent Networks (Redes Recurrentes Simples)
LTSM	Long-Term Short Memory (Red de Memoria Corta a Largo Plazo)
GRU	Gated recurrent Unit (Unidades Recurrentes Cerradas)
ReLu	Rectified Linear Unit (Unidad Lineal Rectificada)
MSE	Mean Square Error (Error Medio Cuadrático)

Operadores usados en redes neuronales:

x_i	Entrada en un tiempo t
s_t	Estado oculto
o_t	Salida en un tiempo t
U, V y W	Matrices de pesos sinópticos
c_{t-1}	Información guardada en la memoria por la etapa anterior
c_t	Información guardada que es pasada a la siguiente etapa
ϕ	Función de activación
$[[]]$	Concatenación de entradas
\oplus	Suma de tensores
σ	Función sigmoide
$x_{i \times n}$	Entrada en un tiempo t para un piso en específico
$w_{i \times j}$	Peso sinóptico de una neurona
$h_i(t)$ ó $net_i(t)$	Regla de propagación
$a_i(t)$	Estado de activación actual
$a_i(t-1)$	Estado de activación anterior
$f_i(t)$	Función de activación
ξ_i	Umbral de la función de activación
$y_i(t)$	Función de salida en un piso i
m	Número de capas que tiene la red neuronal
W^m	Matriz de pesos sinópticos de una capa m
b_i^m	Pesos sinópticos de las conexiones entre las neuronas en la capa m
b^m	Matriz de pesos sinópticos de las conexiones entre las neuronas en la capa m
a^{m+1}	Salida de la neurona dentro de las capas ocultas
$y(k)$	Salida deseada en una interacción k
$\hat{y}(k)$	Salida estimada en una interacción k
e	Error medio cuadrático en una interacción
E	Vector de error medio cuadrático
s_i^m	Sensibilidad ante un cambio en la entrada i
s^m	Vector de sensibilidad del algoritmo de retropropagación

Operadores usados en el modelo matemáticos de la estructura civil:

$x(t)$	Posición en un tiempo t
$v(t)$	Velocidad en un tiempo t
$a(t)$	Aceleración en un tiempo t
A	Amplitud
ω	Frecuencia excitante o forzada ($\frac{rad}{sec}$)
t	Instante en un tiempo (s)
ϕ	Ángulo de fase
T	Periodo (s)
f	Frecuencia $\frac{1ciclo}{s}$
n	Número de ciclos completos
d_t	Tiempo de muestreo (s)
ω_n	Frecuencia natural (sin amortiguar) ($\frac{rad}{s}$)

ω_e	Frecuencia de excitación externa
u	Desplazamiento; desplazamiento relativo al suelo
u_g	Desplazamiento del suelo (o soporte)
u_{go}	Desplazamiento máximo del suelo
R	Factor de modificación en la fuerza
\ddot{u}_g	Aceleración del suelo (o soporte)
R_d	Factor de respuesta de deformación (o desplazamiento)
g	Aceleración gravitacional
F_e	Fuerzas externas ejercidas en la estructura por un sismo
F_{e_i}	Fuerza externa ejercida que afecta a cada piso en la estructura
x_i	Coordenadas generalizadas
\dot{x}_i	Velocidad en el piso i
\ddot{x}_i	Aceleración en el piso i
R	Función de disipación de Raleigh
L	Función lineal lagrangiano
E_I	Energía total
T	Energía cinética
U	Energía potencial
E_S	Energía de deformación elástica
E_D	Energía disipada por el amortiguamiento viscoso de la estructura
E_H	Energía por deformación de la estructura en el rango inelástico
m_i	Masa en el piso i
c_i	Rigidez en el piso i
k_i	Amortiguamiento en el piso i
M_s	Matriz de masa
C_s	Matriz de rigidez
K_s	Matriz de amortiguamiento
$x(t)$	Vector de desplazamiento relativo con respecto a la posición inicial
$\dot{x}(t)$	Vector de velocidad relativa con respecto a la posición inicial
$\ddot{x}(t)$	Vector de aceleración relativa con respecto a la posición inicial
1	vector que distribuye la señal escalar \ddot{x}_g a cada piso de la estructura
$z(t)$	Variables de estado
$y(t)$	Salidas del sistema en términos de desplazamiento que se presenta en cada piso
z	Fuerza de histéresis
A, β y γ	Regulan el tamaño del lazo de histéresis en el modelo de Bouc-Wen
r	suavidad en las transiciones del lazo en el modelo de Bouc-Wen
α	Razón es la relación de post-rendimiento a pre-ceder (elástico)
en el	modelo de Bouc-Wen
v	Degradación de la resistencia en el modelo de Bouc-Wen
η	Degradación de la rigidez en el modelo de Bouc-Wen

Capítulo 1

Introducción

Considerando que, las estructuras civiles son susceptibles a presentar daño estructural a lo largo de su vida útil debido a la presencia de diversos factores físicos o naturales como: el deterioro por uso, envejecimiento, corrosión, fatiga, o por fenómenos naturales (sismos, vientos fuertes) que inciden de manera negativa en su integridad estructural, entre otros. Por lo que, si se desea conocer el estado de salud en el que se encuentra una estructura civil, se requiere monitorear las propiedades dinámicas del sistema [1]:

- El período de vibración: Es el tiempo que le toma a la estructura civil oscilar de un lado a otro cuando hay presencia de vibraciones.
- Razón de amortiguamiento: Es una medida adimensional que posee una estructura civil para la disipación de energía producto de la fricción interna de los materiales.
- Formas modales: Una estructura civil tiende a vibrar a ciertas frecuencias, la cuales son conocidas como frecuencias naturales o de resonantes. Y a cada frecuencia es asociada ha una forma modal, y ante la presencia de actividad sísmica tiende a adoptar su forma modal al vibrar a esa frecuencia.

Una manera de obtener esta información es a partir de la implementación de instrumentos de medición, de acuerdo al tipo de inmueble; como los sensores, que son capaces de medir aceleraciones, velocidades y desplazamientos. O en su defecto, se debe contar con algoritmos que permitan estimar estas señales para correlacionarlas con el daño potencial que pudieran sufrir la estructura civil.

Es decir, a partir de propiedades dinámicas del sistema y su correlación con el modelo matemático de las estructuras de tipo edificio, podrían ser de utilidad como un indicador sobre la condición estructural actual. Esta tarea es comúnmente conocida como monitoreo de salud estructural (ver figura 1.1 [1]), la cual permite detectar problemas de vibración causados por propiedades mecánicas, geométricas e incluso estéticas.

Por lo que se requiere de una inspección monitoreada para la localización del daño estructural y así, poder realizar un diagnóstico de la salud estructural o implementar una estrategia para conservar la integridad de la estructura civil.

En consecuencia, se requiere de un diagnóstico en tiempo real del estado estructural, mediante la utilización de algoritmos que permitan detectar fallas estructurales.



Figura 1.1: Monitoreo de salud estructural.

1.1. Planteamiento del problema

El diagnóstico de daño en estructuras civiles es un proceso esencial para la evaluación y reparación de edificaciones y otras infraestructuras. Este procedimiento permite identificar y evaluar las posibles fallas y características anormales en la estructura, con el objetivo de determinar la causa del daño y proponer soluciones adecuadas. Lo cual reducirá el mantenimiento correctivo y aumentará las acciones mantenimiento preventivo, asegurando la integridad de la estructura como la seguridad de los usuarios. Así mismo:

- Se requiere un algoritmo capaz de realizar un análisis a partir de las características de la estructura (materiales utilizados, diseño y condiciones de uso). Es importante recopilar toda la información relevante para tener una visión completa del problema y poder tomar decisiones a partir de cambios ambientales, es decir, sea capaz de poder adaptarse a las condiciones climáticas que puedan afectar la rigidez de la estructura.

Por lo que es necesario obtener los datos necesarios, o de lo contrario se requiere realizar una inspección visual de la estructura. Donde se examinan los elementos clave como muros, columnas, vigas, losas y cimientos, buscando signos evidentes de deterioro, como grietas, desplomes o corrosión.

Además, se pueden utilizar técnicas no destructivas, como la tomografía de resistividad eléctrica o el análisis ultrasonido, para obtener información más precisa sobre posibles daños ocultos, pero siendo métodos en los que se requieren equipos costosos y de personal capacitado para operarlo.

En este sentido, se presenta la oportunidad de desarrollar un algoritmo para el diagnóstico de daño, que presente la capacidad de detectar, localizar y pronosticar daños en estructuras civiles, pese a recibir perturbaciones o ruidos. Por lo que, se necesitará de mediciones, evaluación de las cargas y deformaciones que presente la estructura.

Buscando así que, el algoritmo sea capaz de lograr un diagnóstico de daño sin la necesidad de modificar el modelo empleado originalmente, ya sea ante la presencia de fuertes lluvias, el terreno incrementa su rigidez o en el caso de que una estructura presente fisuras.

De este modo, se puede explorar la implementación de redes neuronales para detección de

daños. Permitiendo estudiar técnicas de aprendizaje automático a partir de los registros obtenidos, a través del uso de instrumentos de medición especializados, como sensores de carga, extensómetros y acelerómetros, que permiten obtener datos objetivos sobre el comportamiento de la estructura bajo diferentes condiciones. De esta manera, se puede determinar si existen desviaciones significativas respecto a las condiciones de diseño o no.

1.2. Objetivo general

Diseñar un algoritmo basado en redes neuronales para el diagnóstico de daños en estructuras de tipo edificio que proporcione una oportuna localización del piso dañado en el menor tiempo posible, permitiendo una adecuada toma de decisiones o evacuación en el peor de los escenarios. Para reducir de forma favorable las pérdidas materiales y/o humanas.

Por esta razón, los objetivos específicos son:

- Desarrollar un algoritmo de identificación implementando Redes Neuronales para detección de daño.

- Implementar un modelo de histéresis para detectar la gravedad del daño.

- Comparar el algoritmo propuesto mediante índices de desempeño para evaluar la eficiencia.

1.3. Justificación

El territorio nacional (México), presenta una variedad de fenómenos meteorológicos, los cuales, se dan en la atmósfera y de acuerdo a su intensidad, pueden tener efectos positivos o negativos en los ecosistemas, pero que afectan directamente a la salud estructural de los edificios, provocando: agrietamiento, desplazamientos, ablandamiento del asfalto, dilatación de los materiales metálicos que causan fallos estructurales [2].

Así mismo, se presentan: sismos y terremotos, siendo movimientos de la corteza terrestre, ya que, la ubicación geográfica del país se encuentra en:

- Zona sísmica (ver figura 1.2 [3]), por el hecho de que se encuentra dentro del área conocida como el Cinturón de Fuego del Pacífico, donde se concentra la mayor actividad sísmica del planeta. Debido a que, se encuentra situado entre cinco placas tectónicas: Caribe, Pacífico, Norteamérica, Rivera y Cocos [4].

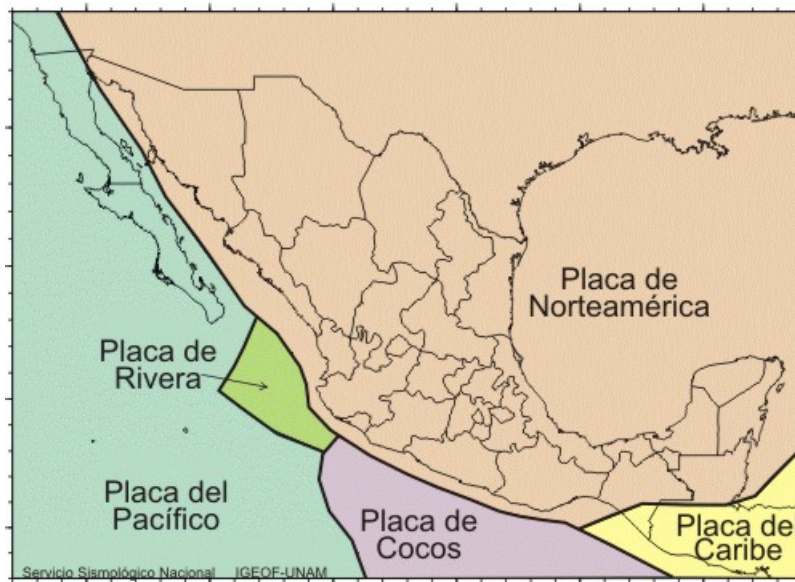


Figura 1.2: Placas tectónicas que afectan a la República Mexicana.

Siendo esta la principal motivación para realizar el presente proyecto, debido a que, México es considerado un país con actividad sísmica frecuente, en el cual se presentan movimientos telúricos de baja intensidad todos los días, y al menos se registra uno de magnitud moderada al año. Siendo estos lo que, provocan mayor daño en las estructuras.

Por esta razón, evaluar algoritmos que permitan una oportuna detección, localización y pronóstico de daños en estructuras civiles, que facilite el analizar los datos pronosticados para obtener un diagnóstico, y en caso de no ser favorable, se puedan tomar las mejores decisiones de mantenimiento, restructuración o demolición, según sea el caso.

De modo que, se obtenga un monitoreo continuo, y si se habla de una estructura civil con un número significativo de pisos, esto proporcionará la información oportuna para la toma de decisiones a partir de una evaluación de los datos recabados.

1.4. Estructura de la tesis

En el capítulo dos se describen los métodos existentes para el diagnóstico de daño en estructuras de tipo edificio, exponiendo las ventajas y limitaciones de cada uno de ellos. Así como un panorama sobre las redes neuronales donde se denota: clasificación, comparación y aplicaciones.

Posteriormente, se presenta el capítulo tres, que consta del modelo matemático de las estructuras de tipo edificio. Adicionalmente, se presentan conceptos fundamentales y el modelo de histéresis a utilizar que permita realizar un diagnóstico de daño.

Después, en el capítulo cuatro, se describe el desarrollo del algoritmo de diagnóstico de daño basado en redes neuronales, así mismo una comparación de dos algoritmos metaheurísticos presentados en el capítulo dos a partir del empleo de simulación e índices de daño.

Finalmente, en el quinto capítulo se presentan las conclusiones a partir del análisis de los resultados obtenidos, dando respuesta a los objetivos e hipótesis planteadas. Además, se presentan algunas sugerencias para investigaciones futuras.

1.5. Aportaciones

Desarrollar un algoritmo de identificación de daños estructurales, el cual será evaluado en un prototipo de edificio de 2 niveles localizado en el Laboratorio de usos experimentales en el Departamento de Control Automático del Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional Unidad Zacatenco.

Esto brindará la oportunidad de implementar el algoritmo en prototipos más grandes o llevarlo a modelos reales, para mantener un diagnóstico constante del estado actual e integridad de los inmuebles.

Por lo cual, el proyecto se vuelve de vital importancia para tener un seguimiento de la salud estructural en tiempo real y poder así tomar medidas de mantenimiento preventivo en caso de ser requeridos.

Capítulo 2

Estado del arte

2.1. ¿Qué es el diagnóstico estructural?

Un diagnóstico estructural es un estudio técnico cuyo objetivo es la verificación de la estabilidad y resistencia mecánica de una edificación, donde se cumplan con los requisitos exigidos por la normativa vigente de seguridad [5].

Un aspecto importante del diagnóstico mencionado son los parámetros que proporcionan la estabilidad de la estructura, los cuales al ser modificados son considerados como causantes de daños en la estructura [6] y estos se presentan por:

- Procesos mecánicos: Deformaciones, grietas, fisuras, desprendimiento parcial en superficies y erosión mecánica por desgaste de materiales.
- Procesos físicos: Se consideran los factores meteorológicos como: la temperatura, el viento y la lluvia, que afectan generando humedad y erosión en la superficie de los materiales y deslizamiento o derrumbe de las estructuras.
- Geológicos: La localización de la estructura, se encuentra cerca o edificado en una falla tectónica.

En los diagnósticos estructurales se emplean métodos [7] que permiten realizar estudios en todo el edificio o centrarse en un piso específico de la edificación, como son:

- Métodos físicos: Magnetometría, fotogrametría, termografía, fotografía ultravioleta, métodos acústicos, medición empleando ultrasonido. Con estos métodos se emplean equipos físicos que requieren ser trasladados a la ubicación de la edificación, los cuales en diferentes escenarios de daños no permiten que los análisis sean llevados con la profundidad necesaria o en todas las partes de la edificación.
- Métodos analíticos: Basados en frecuencia de vibraciones, análisis modal, propagación de onda y en identificación empleando redes neuronales, las cuales se abordaran más adelante en este capítulo.

Los diagnósticos permiten la evaluación y/o estudio de los daños existentes para la realización de los cálculos estructurales necesarios para el mantenimiento preventivo y/o correctivo del inmueble, y en el peor de los casos, la demolición del mismo.

2.2. Tipos de diagnósticos estructurales

2.2.1. Rehabilitación de edificios

Para ello, en el mismo diagnóstico se hace una evaluación en la seguridad [8] como:

- **Análisis histórico:** Evaluación del comportamiento estructural desde su construcción, el cual permite diagnosticar comportamientos futuros, indicando el nivel de seguridad con el que cuenta la estructura actualmente.
- **Análisis cualitativo:** También llamado como procedimiento inductivo, se realiza mediante la comprobación del estado actual de la estructura en la que se lleva a cabo el estudio; comparándola con edificaciones de similar edad estructural que hayan sido evaluadas y confirmado su estado.
- **Análisis cuantitativo:** Se consideran los métodos basados en la ingeniería estructural, obteniendo así, un modelo matemático de las estructuras de tipo edifico.

2.2.2. Adquisición de datos

Para todo estudio existe la adquisición de datos [9], y en los estudios estructurales se debe tener una base de datos existente y confiable.

- **Reunir la información histórica:** Se reúne la información de la estructura desde su año de edificación hasta la actualidad, la cual consiste en las técnicas empleadas en la construcción, así como de las modificaciones realizadas y su desempeño al pasar de los años, donde se considera también la situación actual de toda la edificación.
- **Determinación de características estructurales:** Las propiedades geométricas y mecánicas, se consideran los planos estructurales, el levantamiento de daños y deterioros a lo largo de su vida estructural. Donde se determina si los daños se han estabilizado.
- **Estudios de campo:** Realización de pruebas que determinan las propiedades mecánicas de los materiales estructurales y del subsuelo donde el inmueble está construido.
- **Monitoreo de daños:** Se consideran principalmente las grietas visibles o no visibles, desplomes de cemento o de partes que sean de soporte del edificio y asentamientos.

2.3. Métodos de identificación

La vigilancia o monitoreo de la salud estructural, conocido en inglés como; Structural Health Monitoring (SHM), implica la observación y el análisis de un componente o sistema a lo largo del tiempo, para identificar la variación de alguna de sus características (físicas, químicas o eléctricas), en los términos más generales; daño se define como cambios introducidos en un sistema que afectan adversamente su desempeño actual o futuro.

Siendo un concepto implícito en esta definición está el concepto de que el daño no es significativo sin una comparación entre dos estados diferentes del sistema, uno de los cuales se supone que representa el estado inicial y, a menudo, sin daños, por lo tanto, la definición de daño se limitara a los cambios que se presentan en las propiedades materiales y/o geométricas de estos sistemas, incluidos los cambios en las condiciones de entorno y de comunicación del sistema.

En diversas infraestructuras de ingeniería como la civil, mecánica, entre otras, se llega a conocer como monitoreo de la salud estructural; el cual es un proceso donde se implica la observación de una estructura o de un sistema mecánico a lo largo de un tiempo; donde es utilizado un sistema de mediciones espaciadas periódicamente para una extracción de características de sensibilidad al daño de estas mediciones y se logra hacer un análisis estadístico de dichas características, para la determinación del estado actual de la salud del sistema.

Algunos de los métodos más comunes se basan en los cambios que presenta la estructura civil ante la presencia de fuerzas externa, las cuales generan curvas de comportamiento de las formas modales, energías de deformación, rigideces y flexibilidades, las cuales son medidas dinámicamente.

- Método del cambio de la matriz de flexibilidad, medida dinámicamente (FM Flexibility Method).
- Método del cambio de la matriz de rigidez, medida dinámicamente (SM - Stiffness Method).
- Comparación de curvaturas de las formas modales mediante Enfoque Mixto (MA – Mixed Approach).

Métodos con enfoque de Transformada de la Ondaleta (Wavelet).

- Método basado en la Transformada Continua Wavelet (CWT - Continuous Wavelet Transform).
- Método Mixto.

2.4. Métodos basados en frecuencia de vibración

La comunidad de ingeniería civil se ha dedicado al estudio de la evaluación de daños basada en vibraciones de estructuras de puentes desde principios de la década de los años 80's, haciendo que la investigación en la identificación de daños basada en vibraciones se haya expandido rápidamente en el último par de años; donde la idea básica detrás de esta tecnología es, que a partir de los parámetros modales definen el comportamiento de las propiedades físicas de la estructura.

Por lo que, el análisis de vibración se define como un estudio en el cual se requiere medir los niveles de vibración y su frecuencia natural de vibración del sistema para analizar los datos obtenidos con la finalidad de monitorear y detectar daños que se presentan en las estructuras civiles. Donde la frecuencia natural asociada a su vibración está relacionada con sus propiedades físicas (densidad) y mecánicas (módulo de elasticidad). En consecuencia, el cálculo de frecuencias naturales y formas modales se denomina análisis modal, de frecuencia y de modo normal.

Por ende, este tipo de método requiere de la obtención de modos normales que corresponden a las frecuencias en las que cuando vibra cada piso de la estructura cuando vibran a la misma frecuencia, pero manteniendo una diferencia de fase constante, así poder obtener el periodo fundamental de vibración en una estructura. De manera que se realiza mediante estimaciones de los desplazamientos que sufre la estructura, o a partir de instrumentos de medición de vibración como: acelerómetros, transductores de velocidad o sondas de desplazamiento.

2.4.1. Trabajos hasta el momento

A continuación se presentan trabajos realizados, con el enfoque de frecuencias de vibración:

Creación de bases de datos

Se crea una base de datos a partir de mediciones disponibles, las cuales permiten identificar tipos de daños específicos a partir de características particulares de la forma de las vibraciones; donde generalmente se conoce la ubicación aproximada del daño, lo que es hecho por un analizador de transformada rápida de Fourier (FFT).

Por lo que, presenta una variabilidad de resultados que depende de las condiciones ambientales o de operación, siendo esto un problema, por lo que se requiere la normalización de datos de forma temporal para facilitar la comparación, así como un proceso de limpieza de datos por las personas directamente involucradas en la adquisición de datos. Y se debe tener en cuenta, que la parte de adquisición y limpieza de datos de un proceso de monitoreo de la salud estructural, no debe ser estática [10].

Identificación por medio de actualización de matrices de rigidez

Se basa en la solución de una serie de Taylor de primer orden, que minimiza una función de error de las perturbaciones de la matriz. Mejor conocida como método de actualización basado en la sensibilidad. Siendo el vector de perturbación de parámetros calculada por un método interactivo como Newton-Raphson, el cual permite la minimización de la función de error, pero al ser interactivo se presenta un error por diferencia principal debido al número de decimales pedidos hasta obtener el error relativo sea aproximadamente menor al error fijado [11].

Cambios en características de vibración

Sin importar la base de datos que se emplee, se presentaran errores de sesgo. Así mismo, los valores de las frecuencias finitas que se obtienen son a partir de las condiciones ambientales cambiantes durante la prueba, provocando que los parámetros modales identificados sean menos representativos de las verdaderas propiedades dinámicas de la estructura. Ya que cuando se trata de estudiar sistemas que cambian de lineal a no lineal, las series de tiempo por sí solas (en realidad, su frecuencia, espectros de potencia de dominio) son suficientes para identificar daños [12].

Evolución no Destructiva, conocido en inglés como Non-destructive evolution (NDE)

La implementación de NDE es utilizada principalmente para la caracterización y verificación de la gravedad de los daños, cuando se conoce la localización de estos. Por otra parte, cuando se usa el enfoque de control estadístico de procesos, conocido en inglés como Statistical Processes Control (SPC), se basa en el proceso en lugar de estructura y utiliza una variedad de sensores para monitorear los cambios que se presentan en el proceso. Y así poder obtener las causas que generan que una estructura de tipo edificio presente daño. Y una vez que se ha detectado el daño, se utiliza un pronóstico de daño, conocido en inglés como Damage Forecast (DP) para determinar la vida útil de la estructura para su uso, pero se presenta el problema del pronóstico de daños, ya que no puede predecir dónde o

cuándo se presentará un daño de severidad tal, que la estructura se vea comprometida por completo [13].

2.4.2. Ventajas, limitaciones y desventajas

Ventajas

Al tratarse de una metodología no invasiva, presenta ventajas [14] como lo son:

- Se puede realizar un estudio general de la salud estructural, aún cuando éste se encuentre en lugares de difícil acceso o instrumentación.
- Se puede realizar aún cuando la instalación se presente en un estado de servicio.
- Las mediciones necesarias se pueden realizar con un número mínimo de instrumentación de medición, sin que la precisión en la localización de daño se vea afectada.

Limitaciones y desventajas

Por otra parte, las limitaciones y desventajas de esta metodología son:

- Al utilizar métodos basados en la transformada Wavelet, se tiene que en casos de presencia de daño severo detectan con precisión la localización del daño, pero en caso de daños ligeros no es posible distinguir los errores causados por el método a los datos obtenidos en la presencia de daño.
- Al implementar un enfoque mixto, la elección de los factores de limitación jugarán un papel importante, ya que determinarán la precisión, con la que se localice el daño en la estructura.
- Los errores causados por presencia de ruido en las mediciones se intensifican cuando no se realiza una filtración adecuada de la señal.

2.5. Métodos basados en análisis modal

La mayoría de los métodos para el monitoreo de la salud en ingeniería estructural, civil y mecánica se basan en el análisis modal. El cual es un método experimental, donde se define el modelo dinámico de una estructura a partir de determinar de las propiedades dinámicas o parámetros modales, los cuales son:

- Frecuencia de vibración.
- Amortiguamiento de la estructura.
- Modos de vibración.

Las propiedades dinámicas dependen de la masa y rigidez que presente la estructura, por lo que es necesario identificar las frecuencias de vibración y conocer la respuesta de la estructura cuando una fuerza actúa sobre la misma. Lo que permitirá conocer la deformación provocada en la estructura compuesta por una combinación o una mezcla de todos los modos de vibración.

Por lo que, se emplea el análisis modal para determinar los modos de vibración, siendo estos los patrones o forma característica del sistema al momento de vibrar [15].

2.5.1. Trabajos hasta el momento

A continuación se presentan trabajos realizados, con el enfoque de análisis modal:

Cálculo de formas modales y periodos de vibración para edificios altos, empleando modelos continuos

Se considera el modelo flexión-cortante en paralelo, modelo en el cual se presenta una viga en flexión con ISE (Interacción Suelo Estructura), obteniendo como resultados el modelo de un edificio tubo en tubo, que consiste en el entramado de vigas y columnas con nudos empotrados entre sí.

Verificando así, la equivalencia de los parámetros modales que definen la variación de la forma modal con respecto a la altura entre el modelo de un edificio tubo en tubo, de una forma rápida. Por lo que, se queda limitado a modelos en los cuales se considere una base empotrada y principalmente estructuras esbeltas [16].

Identificación de propiedades dinámicas empleando modelo de una estructura

En este caso se presenta un trabajo el cual se desarrolló un prototipo experimental de una estructura, la cual fue sometida a pruebas de vibración ambiental (viento y temperatura), así como de vibración forzada utilizando una mesa vibradora para generar la simulación de actividad sísmica y obtener así el comportamiento de la estructura.

Lo que permitió obtener las frecuencias de vibración, formas modales y estimaciones del amortiguamiento crítico empleando análisis espectral. Por lo que, si se desea tener mejores resultados, se requiere emplear más instrumentos de medición que permitan registrar de forma simultánea las vibraciones en diferentes puntos de la estructura, así como el empleo de un método de corrección por sobreestimación, debido a que el mecanismo empleado al tener diferente número de entradas, produce una sobre apreciación de los resultados obtenidos [17].

Método de diagnóstico local empleando metodología no probabilística y método de series de tiempo

La integración del modelo de reconocimiento de modos de vibración y un modelo autorregresivo de medio móvil, conocido en inglés como AutoRegressive Moving Average Models (ARMA), permite detectar daños significativos, como no significativos, así mismo con y sin presencia de perturbaciones ambientales. Pero necesita una cantidad grande de instrumentos de medición (acelerómetros) y realización de cálculos complejos, así mismo presenta problema en la exactitud para localizar el daño [18].

Análisis de vibración a partir de mediciones en diferentes casos de estudio

Se analizan como el puente del Cañón de Alamosa que es el caso de estudio, presenta diferentes frecuencias naturales y formas modales. Ante la presencia de diversos factores como lo son: la cantidad de tráfico y las condiciones ambientales, implementando el uso de cambios en los parámetros modales para identificar la ubicación y el alcance del daño en las estructuras [19]. Por lo que, se necesita obtener mediciones de las propiedades dinámicas de estructura que se desea estudiar en diferentes épocas del año, diferentes condiciones climáticas y en condiciones de servicio, por lo que al emplear el análisis modal se presentan errores aleatorios introducidos durante el proceso de medición y selección de datos. Haciendo que la localización de daño estructural cause cambios en los parámetros si el daño localizado

se encuentra fuera de estos límites permitidos para que realizar una exacta localización de daño [20].

Análisis de vibración usando series de tiempo

El enfoque se basa en la premisa de que si hubiera daño en la estructura, el modelo de predicción previamente identificado utiliza la serie de tiempo cuando la estructura no está dañada, no podría reproducir la serie de tiempo obtenida cuando la estructura se encuentra dañada. Además, que el aumento de errores residuales se maximizaría en los instrumentos de medición, si se encuentran cercanos a las ubicaciones reales de los daños estructurales. Por lo cual, el método no ha demostrado ser efectivo para detectar daños en un estado temprano [21].

2.5.2. Ventajas, limitaciones y desventajas

Ventajas

El análisis modal es la opción más utilizada para obtener el comportamiento de las estructuras ante la presencia de actividad sísmica, cuando estas se encuentran construidas de concreto reforzado o modelo de tubo en tubo.

Limitaciones y desventajas

Mientras que las limitaciones y desventajas de esta metodología son:

- Se requiere realizar un mayor número de experimentos aumentando el tiempo de ensayo, por ende aumentando el financiamiento económico requerido para llevar a cabo el estudio [22].
- Los trabajos presentados, no se han aplicado en modelos de estructuras reales [18].
- Requiere de un monitoreo continuo mediante instrumentos de medición que permitan adquirir los parámetros o la filtración necesaria para eliminar las molestias ambientales.
- Se presentan problemas en la exactitud para la localización de daño estructural.

2.6. Métodos basados en propagación de onda

Debido a que las frecuencias naturales de un sistema estructural presentan variaciones no lineales, bajo distintos niveles de excitación, siendo atribuidas al tipo de suelo en el que se encuentra situada la estructura; en diferentes trabajos escritos, el ablandamiento estructural no lineal no está bien caracterizado o definido.

De modo que, se debe tenerse en cuenta si se va a utilizar de manera efectiva el monitoreo del edificio en tiempo real como una herramienta de evaluación de daños posteriores a la actividad sísmica. Es decir, los cambios observados en las frecuencias fundamentales se correlacionan con el clima, el historial de terremotos y el uso de los edificios [23].

Por lo cual, dentro de los métodos que evalúan las vibraciones, se estudia el análisis de propagación de ondas en la estructura, siendo un método que permite conocer el daño causado por la actividad sísmica, a partir del análisis de propagación de las ondas flexionantes. La

cual permite determinar cómo se distribuyen las fuerzas internas en una estructura encontrando la fuerza cortante y el momento flexionante, siendo determinantes para garantizar la resistencia y estabilidad en una estructura [24].

2.6.1. Trabajos hasta el momento

A continuación se presentan trabajos realizados, con este enfoque de propagación de ondas:

Identificación de velocidad de onda de corte

Se plantea la identificación paramétrica de una estructura, utilizando un modelo de barra elástica y discretizada, que cuenta con un punto de corte. En la cual se emplea una identificación por mínimos cuadrados normalizados con factor de olvido y un esquema de proyección paramétrica que asegure la convergencia, para así obtener la velocidad de onda de corte, la cual permite calcular la rigidez de la estructura de los pisos o el módulo cortante del suelo en el que se encuentra.

Por lo que implementando este tipo de identificación permite resolver el problema que se presenta en las metodologías anteriores en las que se requiere desacoplar los parámetros rigidez-masa y amortiguamiento-masa. Así mismo, elimina perturbaciones constantes y atenúan el ruido de medición al utilizar una parametrización filtros lineales de tipo integral, pero requiere de señales de aceleración medibles [25].

Identificación de rigidez en los entrepisos de un edificio

Se plantea una detección de daños por identificación de rigidez en los entrepisos, los cuales separan horizontalmente un piso de otro dentro de los diferentes niveles de una estructura. Por cuál esta identificación se basa en la programación de un pulso de Ricker a través de la estructura, generando tiempos de arribo en cada piso.

Por lo tanto, se puede replantear la ecuación de rigidez en función de los retardos de tiempo obtenidos. Y para determinar el tiempo propagación de onda son calculados aplicando el método de Haskell, permitiendo así que se pueda usar este enfoque a edificios de múltiples pisos [26].

2.6.2. Ventajas, limitaciones y desventajas

Ventajas

Se trata de una metodología fácil y rápida que evalúa la integridad y el funcionamiento de los edificios, luego de estar sometidos a activa sísmica de gran intensidad. Así mismo, presenta fácil una adaptación a los instrumentos de medición disponibles.

Limitaciones y desventajas

Mientras que la limitación de este enfoque es debido a que se desarrolla en el dominio de la frecuencia, por lo que no se puede emplear en tiempo real, así mismo la precisión para identificación de los parámetros y en el tiempo requerido para la estimación son fundamentales para la correcta detección de daño.

2.7. Método basado en redes neuronales

El método de redes neuronales está enfocado a evaluar el daño causado por actividad sísmica en cada piso de una estructura. De modo que, para utilizar una red neuronal se emplean dos bases de datos, la primera contara con datos de la estructura, la cual no presentará daño, y la segunda será cuando la estructura presente daño, para así poder realizar la predicción de daño, a partir del entrenamiento realizado con la primera base de datos obtenidos para el entrenamiento de la red neuronal.

Por consiguiente, se requiere obtener las bases de datos que serán empleadas como entradas para la red neuronal, las cuales se encuentran compuestas por las frecuencias naturales a las que vibra la estructura cuando se presenta actividad sísmica. Y poder así utilizar la base de datos obtenida a partir del método de los elementos finitos para el entrenamiento de la red neuronal.

Mientras que para predicción de daño en cada nivel de la estructura, se obtiene de la comparación de las matrices de rigidez de la estructura previstas por el método y la identificada a partir de pruebas pseudodinámicas, por datos experimentales que proporcionan las características modales de la estructura estudiada.

En consecuencia, el interés en este método de identificación de daño en estructuras y/o sistemas mecánicos ha crecido en las últimas décadas por razones económicas y de seguridad. Lo que ha llevado a desarrollar numerosos métodos basados en vibraciones para detectar el daño en etapas tempranas.

Por otra parte, en la última década se ha explorado un enfoque basado en redes neuronales para la identificación de daño, el cual permite emplear modelos experimentales tridimensionales simplificados de una estructura compuestos por barras y masas concentradas.

De modo que, los modelos experimentales de estructuras pueden ejemplificar los diferentes tipos de rigidez de vigas mixtas y de la densidad del material, como lo es el hormigón, debido a que las variables físicas son más inciertas. Son considerados como parámetros que serán actualizados, de tal forma que se parten de estos para obtener las base de datos de que contiene las frecuencias naturales de la estructura, empleado un modelo flexionante el cual realiza una variación aleatoria de los parámetros ya mencionados.

2.7.1. Trabajos hasta el momento

A continuación se presentan trabajos realizados, con este enfoque de redes neuronales:

Detección de daño empleando una red neuronal multicapa o perceptrón multicapa, conocido en inglés como; Multilayer Perceptron (MLP)

Se emplea una red multicapa, por lo que se requiere recaba una base de datos, la cual se filtra para evitar la generación de errores por ruido de medición que afecte el aprendizaje y obteniendo así una base de datos final conformada por las frecuencias naturales de la estructura estudiada.

En este caso se emplea un algoritmo de aprendizaje supervisado conocido como retropropagación, con la finalidad de reducir el error en la predicción de daño que presente cada viga. Donde el error es definido como el cociente entre la rigidez a la flexión de la viga dañada y la de la no dañada, obteniendo así que la red multicapa diseñada presenta un error en la predicción de daño del 5 y 10% con respecto a la estructura se encuentra sin daño.

Pero si, se considera la rigidez de las columnas y vigas en el modelo, se obtiene la rigidez total de cada piso de la estructura, por el cual el parámetro de daño se puede definir como

la relación entre la rigidez de la viga dañada y la no dañada. Por lo tanto, la matriz de rigidez obtenida para la estructura no dañada es ahora más cercana a la experimental, con diferencias absolutas entre elementos comprendidas entre 0,04 y 0,41 %, llegando a que el modelo reproduce mejor las propiedades físicas de la estructura [27].

Un aspecto importante a considerar en este modelo de red neuronal es que se realiza el aprendizaje y prueba, fuera de línea, esto quiere decir, que se debe generar las bases de datos para así poder utilizarla.

Método Nelder-Mead SimplexRed combinado con una red multicapa

Se diseña una red multicapa combinada con el método Nelder-Mead Simplex para detectar daño, el cual consiste en emplear este método numérico para minimizar la función de objetivo en un espacio multidimensional como lo es la red multicapa.

De modo que este método busca a partir de aproximaciones una solución óptima local con N número de variables, cuando la función a minimizar varía suavemente. Debido a esto, los ingenieros pueden emplearlo en la detección de daño, ya que se puede fijar los valores de entrada y considerar factores externos tales como: vibraciones, vientos, materiales de construcción, entre otros.

Esto se debe a que el método genera una nueva posición en cada interacción a través de extrapolación del comportamiento de la función en los vértices de un simplex. Debido a esto no es necesario calcular todos los valores posibles de la función, ya que el algoritmo reemplaza los puntos de prueba ajustándose para así encontrar la solución que minimiza la función más rápidamente

Por lo cual, no solo se requiere tener la base de datos que contenga las frecuencias naturales de la estructura, sino también la flexibilidad modal de cada piso de la estructura. Por lo tanto, se deberá definir una cantidad específica de modos idénticos permitidos para no generar conflicto con la definición de simplex.

Lo lleva a que el error máximo obtenido en la detección de daño es menor al 0,05 en 80 % de los casos que se estudiaron, por otra parte, se obtuvo una identificación errónea igual 3 %, un error mayor al 0.2 en 2 % en las simulaciones realizadas y solo en el 1 % de las simulaciones no se detectó el daño.

En este caso se presentan ventajas: de basarse en consideraciones geométricas en vez de requerir la derivada de la función objetivo, así mismo en las primeras interacciones se busca un mínimo más amplio para así evitar caer en mínimos locales que dificulten el aprendizaje. Mientras que la desventaja que presenta es que se requiere un mayor número de interacciones aumentando así el tiempo que se requiere para la convergencia [28].

Detección de daño empleando redes neuronales convolucionales para identificación y redes neuronales convolucionales para clasificación, conocidas en inglés como Convolutional Neural Networks for Identification (CNNFI) y Convolutional Neural Networks for Classification (CNNFC) respectivamente

En este caso se diseñan redes neuronales de tipo CNNFI y CNNFC, para la localización de daños en estructuras de difícil acceso para equipos de diagnósticos. Por lo cual para evaluar su desempeño se requiere realizar pruebas más exhaustivas.

Por lo que, se emplea una combinación de estas redes con un método de histéresis para así validar los resultados obtenidos a través de las redes neuronales. Permitiendo entender como la activada sísmica afecta la disipación de energía en las estructuras a través de la generación de ciclos de histéresis que describen el comportamiento de los materiales.

Pero debido a que este estudio se realizó en un escenario controlado para realizar las pruebas de adquisición de datos y poder así obtener índices de rendimiento mejores que otras redes neuronales diseñadas para detección. Esto provoca que se reduzca la veracidad de resultados ante la presencia diferentes mecanismos de daño debido a la caracterización que se requiere para diferentes ciclos de histéresis [29]. Ya que se empleara la metodología de redes neuronales, en la sección siguiente se profundizará en este método para así poder tener un panorama más completo de la estructura de una red neuronal como de su funcionamiento.

2.8. Redes neuronales

El cerebro humano es considerado una computadora notable, capaz de interpretar información imprecisa, adquirida por los sentidos a una alta velocidad. Con la capacidad de distinguir una voz en específico entre una sala ruidosa, identificar una persona entre una multitud o entender mensajes subliminales en discursos; siendo su capacidad más notable el poder aprender requerir de instrucciones explícitas.

Por la eficiencia que presenta el cerebro al realizar diferentes procesos, han inspirado a investigadores, ha desarrollar la teoría de redes neuronales, conocida en inglés como; Neural Networks (NN), la cual emula el funcionamiento de redes neuronales biológicas. Presentando una alternativa a la computación clásica, para problemas, en los cuales los métodos tradicionales no han sido eficientes o han entregan resultados insatisfactorios.

Por lo que, existen diferentes maneras para definir las redes neuronales, entre las cuales se puede tener una descripción amplia y detallada hasta una corta y genérica. Como se presenta a continuación:

Definición 1. *“Sistema de computación hecho por un gran número de elementos simples, elementos de proceso altamente interconectados, los cuales procesan información por medio de su estado dinámico como respuesta a entradas externas” [30].*

Definición 2. *“Redes neuronales artificiales son redes interconectadas masivamente en paralelo de elementos simples (usualmente adaptativos) y con organización jerárquica, las cuales intentan interactuar con los objetos del mundo real del mismo modo que lo hace el sistema nervioso biológico” [31].*

Por lo cual, las redes neuronales [32] son un método de aprendizaje y procesamiento automatizado, inspirado en el funcionamiento del sistema nervioso del ser humano. De modo que se trata de un sistema que se encuentra compuesto por un conjunto de neuronas (nodos) interconectadas entre sí mediante enlaces (bias) colaborando entre sí para producir un estímulo de salida, pareciéndose así al cerebro humano.

Respecto a cada neurona se considera una entrada o salida, que a su vez es entrada de las neuronas en la capa antecesora, donde cada una de las entradas es multiplicada por un peso sinóptico, sumándose así a los resultados parciales y al implementar una función de activación se calcula la salida. Siendo así la composición más simple de una red neuronal.

En resumen, las redes neuronales van más haya de los sistemas de redes interconectados masivamente en paralelo de elementos simples, por lo cual se emplea una organización jerárquica. De esta manera, las redes neuronales intentan resolver problemas complicados, creando así un sistema adaptable que las computadoras emplean para aprender de sus errores y mejorar continuamente, como lo hace el sistema nervioso biológico(ver figura 2.1 [33]). En la tabla 2.1, se presenta la comparación entre los elementos que componen una neurona biológica con respecto a la neurona artificial.

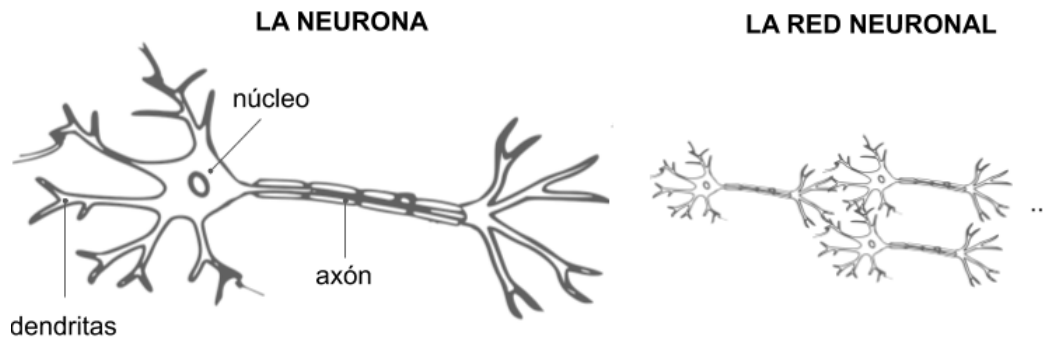


Figura 2.1: Red neuronal biológica.

Neurona biológica	Neurona artificial
Dendritas	Entradas
Sinapsis o Núcleo	Pesos sinópticos
Axón	Salidas

Tabla 2.1: Comparación entre una neurona biológica y una artificial.

2.9. Clasificación de Redes Neuronales

2.9.1. Clasificación según la topología de red

Red neuronal monocapa o Perceptrón Simple, conocido en inglés como; Single Layer Network

La red neuronal monocapa o perceptrón simple [34], es la red neuronal más simple que existe de aprendizaje supervisado. Ya que se encuentra compuesto por una sola capa de neuronas, como se presenta en la figura 2.2. [35]

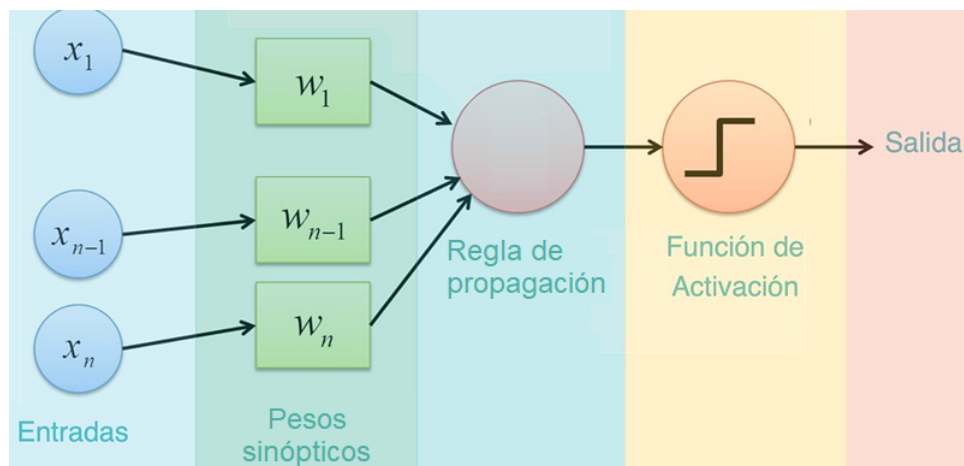


Figura 2.2: Estructura red neuronal monocapa.

El funcionamiento de la red neuronal monocapa, consiste en leer los valores de entrada, para así emplear la regla de propagación seleccionada y el resultado obtenido se introduce en una función de activación que entregara el resultado final. En consecuencia, se puede considerar que el funcionamiento es sencillo.

Por otra parte, el proceso de aprendizaje en una red monocapa consiste en encontrar los pesos sinópticos y el umbral (el grado de inhibición de la neurona), a partir de una selección arbitraria de hiperparámetros, que se irán modificando de acuerdo con los valores deseados y los calculados por la red. Por lo tanto, se trata de un proceso adaptativo que aprende de manera iterativa siguiendo estos pasos:

1. Inicialización de hiperparámetros, en este caso: pesos sinópticos y umbrales.
2. Se crea un bucle el cual terminara cuando, los pesos sinópticos calculados nos entreguen un error que se encuentre dentro del margen de error aceptable:
 - Leer valores de entrada.
 - Calcular error.
 - Evaluar error calculado con el error definido como aceptable.
 - Error obtenido mayor al margen de error aceptable, se actualizan pesos sinópticos según el error calculado.
 - Actualización de pesos sinópticos de entradas.
 - Actualización de umbral.
 - Error calculado menor al margen de error aceptable, se detiene el bucle.

En general, este tipo de red solo se recomienda usar en sistemas lineales, debido a que no dispone de capas ocultas que permitan un mejor aprendizaje y así poder obtener una validación para su implementación en sistemas más complejos. Por lo cual la función de activación más empleada para la red monocapa es la función sigmoide.

Red neuronal convolucional, conocida en inglés como; Convolutional Neural Networks (CNN)

Uno de los problemas clásicos en la inteligencia artificial, se presenta en la identificación de objetos en imágenes. Debido a que se realiza mediante un aprendizaje automático, de modo que se basa en redes neuronales.

Sin embargo, al utilizar redes neuronales se presenta un desempeño adecuado cuando se trata de detectar un elemento en imágenes pequeñas, pero si se desea localizar ese mismo elemento en una imagen de mayor tamaño se vuelve insatisfactorio su desempeño. Dado que, se requiere procesar un mayor número de píxeles, lo que implica un aumento exponencial a las variables de entrada (features) que se analizarán, provocando que el manejo de información se dificulte o no la identificación por redes neuronales no sea una opción.

Esto se soluciona al emplear redes neuronales convolucionales [36], debido a que cada parte de la red se le entrena con la finalidad de realizar un análisis, permitiendo reducir la arquitectura de la red, en consecuencia la fase de aprendizaje será más rápida en realizar. Por lo cual, son aplicadas en:

- Problemas de visión por computador.
- Análisis de imágenes.
- Reconocimiento de imagen.

Una red neuronal convolucional es capaz de detectar características simples como: bordes, líneas, etc. y formar mapas de características hasta detectar el elemento buscado. Como sucedió en el 2012, cuando una “convolutional neural network”, ganó por primera vez la competición de reconocimiento de objetos ImageNet, la cual consiste en clasificar 1000 objetos en conjuntos de más de un millón de imágenes [37].

La estructura habitual de una red neuronal convolucional, es de tipo multicapa conformada por dos diferentes capas. Donde cada una se encarga de realizar operaciones específicas dentro de la red, por lo cual es necesario conocer cómo funciona la capa: convolucional y de reducción alternada, para poder usar al final una capa clasificadora de conexión total, como se puede observar en la figura 2.3 [38].

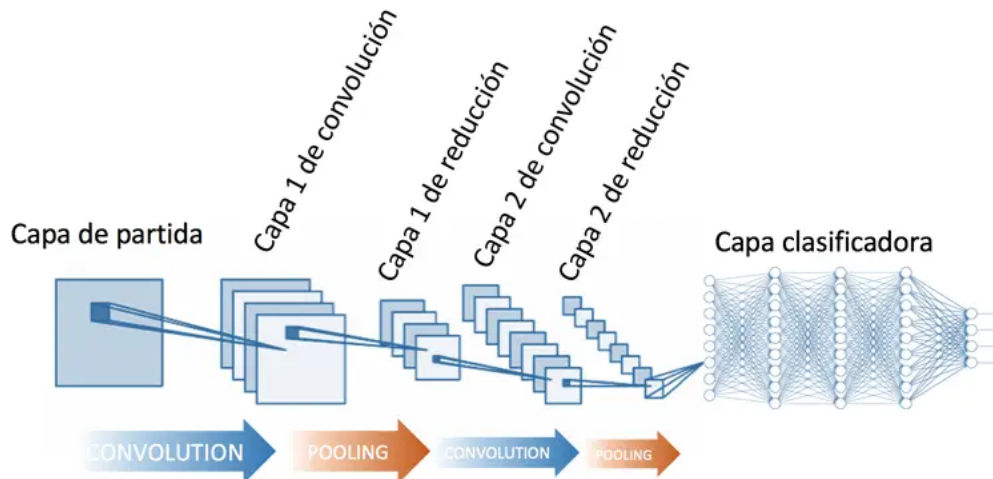


Figura 2.3: Estructura red neuronal convolucional.

Antes de empezar a describir el funcionamiento de las redes neuronales convolucionales, se necesita aclarar que el kernel, en este caso se considera como el filtro que se aplica a la imagen, para poder extraer las características importantes o los patrones que se desean identificar.

En primer lugar, se tiene la capa convolucional, donde se realiza la operación de convolución. La cual consiste en tomar un grupo de píxeles de la imagen de entrada para ir realizando el producto escalar con un filtro (kernel) estos se irá modificando de igual manera que los pesos sinópticos. De modo que se genere un mapa de características (features maps), en el que se encontrará cada posible ubicación del filtro a partir de una imagen de entrada.

Las imágenes que entran, suelen tener dimensiones de $n \times n$ (matriz bidimensional), donde n representa el número de píxeles de alto y ancho. Por otra parte, la definición del número de canales dependerá de los colores de la imagen, ya que si se encuentra en escala de grises, equivaldría a un canal o si se encuentra dentro del espectro visible “RGB” (rojo, verde y amarillo), equivaldría a 3 canales. Por lo cual, es necesario definir y considerar:

- El tamaño de la matriz bidimensional del kernel, es variable y depende del diseñador de la arquitectura, de acuerdo a las necesidades del problema a resolver.
- El número de filtros que se empleen por capa, se determina de igual manera por el diseñador.
- Si la imagen se encuentra dentro del espectro “RGB”, se tendrán 3 filtros del mismo tamaño, los cuales se sumarán para tener una imagen de salida.

- Se requiere normalizar los datos, es decir, se requiere codificar un pixel en 8 bits (1 byte). De modo que, un pixel permite hasta 256 variaciones de color, por lo tanto, cada color tiene un valor específico entre 0 y 255. Por lo cual, para normalizar cada pixel y así encuentre en un rango de valores entre 0 y 1, se dividen los píxeles entre el valor más alto permitido 255.

Se debe de considerar que a mayor cantidad de filtros, la cantidad de características que se podrán extraer será mayor, pero aumentado así el costo computacional. Así mismo, si el tamaño del filtro es grande, el número de pesos sinópticos que se deberán ajustar en la red aumentara proporcionalmente.

La ventaja de utilizar filtros es que su tamaño determinara el número de neuronas que se utilizaran:

Ejemplo 1. Si se tiene un filtro de 4×4 se utilizarán 16 neuronas.

Esta característica sirve para extraer el mismo número de características en cualquier parte de la entrada, así se puede se reduce el número de conexiones y el número de parámetros, que se deben configurar dentro de la capa convolucional, los cuales son:

- Paso (Stride): Establece el número de pasos que dará el filtro mientras recorrer la imagen de entrada para aplicar la operación de convolución como se presenta en la figura 2.4. Es importante que el stride sea de 2 o más, para conseguir que los mapas de características de salida tengan un tamaño reducido en comparación a la entrada.

Ejemplo 2. Recorrido del filtro a través de los píxeles de entrada.

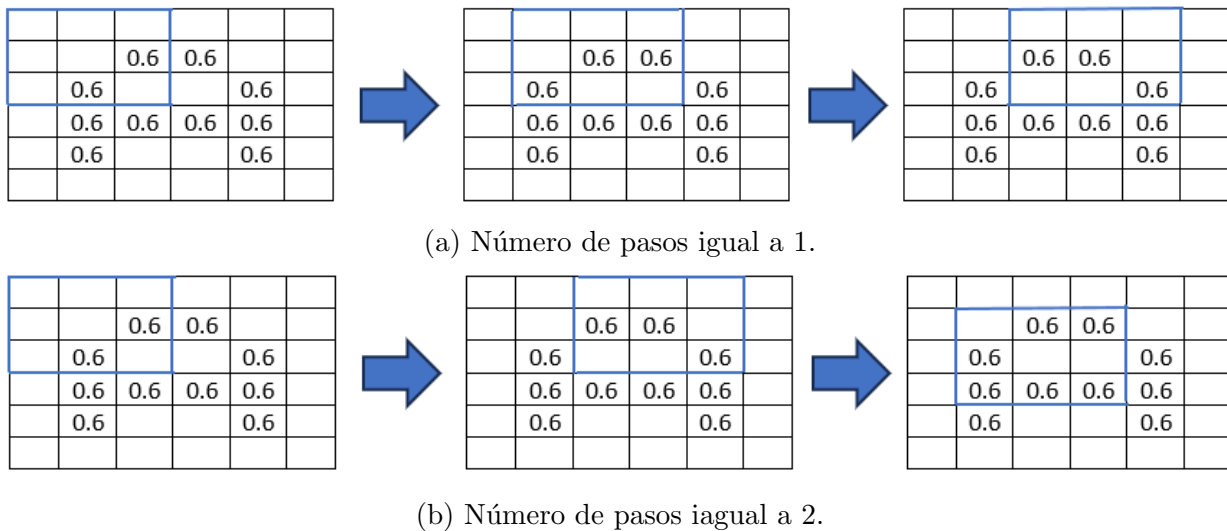


Figura 2.4: Recorrido del filtro de acuerdo al número de pasos definidos.

- Función de activación:** Se aplica después de aplicar la convolución en los mapas de características, para este tipo de redes se utilizan las funciones de activación: tangente hiperbólico o ReLU, debido a que ofrecen una tasa de aprendizaje adecuada y monitorean la cantidad de neuronas muertas, que se presentan cuando la función de activación como su derivada son iguales a cero. También se podrían implementar las funciones de activación: Leaky ReLU o Softmax, pero no se recomienda utilizar la función de sigmoide.

Así mismo, junto a las capas de convolución se añaden otros dos tipos de capa:

- Capa de activación: Se codifica de tal forma, que se encarga de aplicar las funciones de activación a las neuronas, anteriormente vistas, a cada uno de los píxeles de la imagen de salida o a los mapas de características.
- Capa de normalización por lotes (Batch Normalization): Tienen una finalidad de normalizar los mapas de características de salida para que los valores que los compongan estén en un rango normalizado de acuerdo a la función de activación.

En segundo lugar, se emplea la capa de reducción (pooling). Esta capa se emplea después de una capa convolucional o antes de utilizar otra capa convolucional. Tiene la función de disminuir el tamaño de los parámetros al eliminar las características menos repetidas y conservar las características más repetidas, para las siguientes capas convolucionales.

La disminución de parámetros se realiza mediante la extracción de características, que consiste en reducir y codificar las variables sin procesar en características más manejables que describan los de datos originales, Se debe considerar que al realizar esto se pierde precisión, pero se gana compatibilidad. Esta extracción de características se puede realizar mediante: la agrupación promedio (Average Pooling) o el máximo de una región fijada del mapa de características (Max Pooling).

Los parámetros de diseño para este tipo de capa son los stride (saltos), ya que deben ser superiores o iguales a dos, debido a que la función de aplicación son filtros fijos de un tamaño en específico que realizan la aplicación del operador de convolución.

Por lo tanto, se tiene la peculiaridad de que pueden ser o no usadas debido a que si se emplea un stride (paso) de valores superiores o iguales a dos en las capas de convolución se pueden conseguir resultados similares. Es por ello por lo que en algunas arquitecturas de este tipo de red no se utilizan.

Por último, para completar la estructura CNN, se requiere utilizar una capa de salida, la cual es una capa totalmente conectada, conocida en inglés como; Fully Connected (FC) Layers. Ante esta situación, para convertir los mapas de característica en vectores unidimensionales, se emplea una capa que clasifique la salida, de forma tal, que se tendrá una neurona por cada variación de color que se desee predecir.

Algunas de las redes CNN más importantes en los últimos años para reconocimiento de objetos son:

- AlexNet [37]
Se trata de una red compleja para reconocimiento de objetos que las empresas siguen utilizando, y se encuentra compuesta por 5 capas convolucionales, 3 capas de reducción, 3 capas conectadas totalmente al final. Presentando una tasa de error del 25 % al 17 %, al momento de reconocer un objeto determinado.
- VGGNet [39]
Red para reconocimiento de objetos, con arquitectura conformada por 19 capas convolucionales con tamaños de filtro de 3×3 y paso igual a 1, y capas de max pooling con paso igual a 2. Presentando una tasa de error del 7.2 %, al momento de reconocer un objeto determinado.
- GoogLeNet [40]
Red enfocada a reconocimiento de objetos, presenta una arquitectura compleja, ya que es un modelo bastante diferente al general que se ha mencionado. se encuentra

compuesta por varios bloques, en los cuales se realizan diversas operaciones en paralelo, en vez de en forma secuencial. Presentando una tasa de error del 6.7%, al momento de reconocer un objeto determinado.

Red neuronal recurrente, conocida en inglés como; Recurrent Neural Network (RNN)

La idea de las redes neuronales recurrentes [41] surge a partir de querer procesar datos de manera secuencial o temporal, al contrario de las redes:

- Feedforward, en las cuales los datos fluyen en una sola dirección
- Red neuronal convolucional, como se presentó anteriormente, requieren realizar el proceso de convolución y clasificación para obtener una respuesta.

Mientras que las redes neuronales recurrentes, no tiene una estructura de capas definida como se presenta en la figura 2.5 [42], es decir, presentan conexiones arbitrarias entre las neuronas, permitiendo que los datos fluyan de una capa a otra a través de un estado oculto. Lo que permite a la que la red tener memoria, en la cual se almacena información sobre los estados anteriores.

Así mismo, la salida del estado oculto, se actualizan en cada paso de tiempo y en un momento dado se utiliza como entrada para el siguiente paso de tiempo. De manera que, la RNN presenta memoria a largo plazo, siendo capaz de detectar patrones previo a secuencias de datos que se extienden por largos periodos de tiempo.

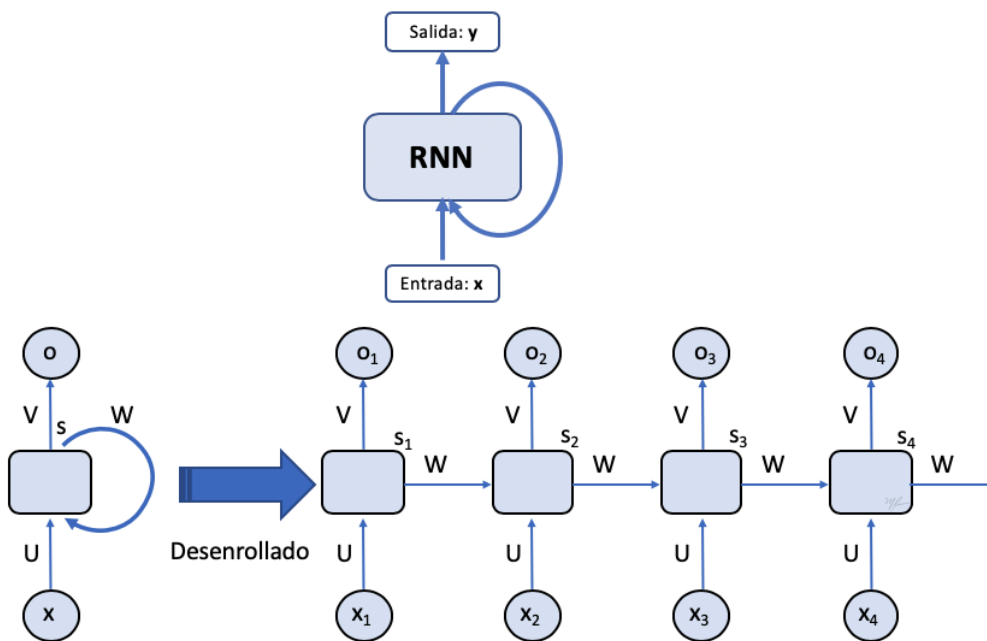


Figura 2.5: Estructura red neuronal recurrente.

Las redes neuronales recurrentes tienen variantes y cada una está sujeta a cierta selección de hiperparámetros, por lo cual se presentaran las más conocidas y los usos de estas.

- Redes recurrentes simples, conocida en inglés como; Simple Recurrent Networks (SRN)
Las redes neuronales recurrentes se incorporan la retroalimentación, lo que se consigue

crear la temporalidad, permitiendo a la red que tenga memoria. Teniendo una estructura que se extenderá igual al número de datos en el vector x , como se presenta en la figura 2.5, se explica de la siguiente forma:

- x_t , es la entrada en un tiempo t .
- s_t , es el estado oculto, ya que no se observa como: dato o salida, en un tiempo t que se calcula como:

$$s_t = f_1(U x_t + W s_{t-1}) \quad (2.1)$$

- o_t , es la salida en un tiempo t , tal que:

$$o_t = f_2(V s_t) \quad (2.2)$$

En este caso la red solo se define como: una sola unidad, celda o cell, pero este procesamiento requiere de tres matrices de pesos sinópticos: U , V y W . Siendo esta la estructura más simple de una red neuronal recurrente.

- Red de memoria corta a largo plazo, conocida en inglés como; Long-Term Short Memory (LTSM)

Dado que las redes neuronales recurrentes, presentan problemas en la fase de aprendizaje, ya que el gradiente de retropropagación es propenso a desvanecerse o crecer desmesuradamente, esto se debe a que no solo depende del error presente sino también de los errores pasados. Haciendo que la acumulación de errores dificulte la capacidad de almacenamiento de la memoria a largo plazo.

Surgen las redes LSTM (Long Short Term Memory) las cuales están compuestas por unidades LSTM como se presenta en la figura 2.6 [42], es un tipo especial de red neuronal recurrente descrito en 1997 por Hochreiter & Schmidhuber [43]. De modo que nos permiten solventar el problema presentado en las RNN, puesto que se incorporan una serie de pasos que permite separar la información que se guardara en memoria de la que se eliminara. La unidad de memoria de término largo, contiene tres puertas que controlan el modo en que la información pasa de c_{t-1} a c_t :

- Puerta de entrada, controla la información que nueva que puede entrar o no.
- Puerta del olvido, se encarga de realizar la selección de información entre datos importantes y superfluos. A partir de la utilización de switches que controlan el “olvido” de la memoria pasada y la nueva información a “recordar”.
- Puerta de salida, es una selección controlada de los resultados obtenidos al procesar el almacenamiento actualizado.

Por lo que son empleadas para: compresión de textos de lenguajes naturales, reconocimiento de escritura a mano, reconocimiento de voz y reconocimiento de gestos, siendo los más notados en los últimos años.

- Unidades Recurrentes Cerradas, conocida en inglés como; Gated recurrent Unit (GRU) Las redes Unidades Recurrentes Cerradas [44] es un tipo especial de red neuronal recurrente, la cual es: unidad de memoria GRU que consta de dos puertas, controlan el flujo de información:

- Puerta de actualización y una puerta de reajuste: Realiza la actualización de datos que se indique de acuerdo con la cantidad de información que se requiera mantener con respecto a las celdas anteriores.

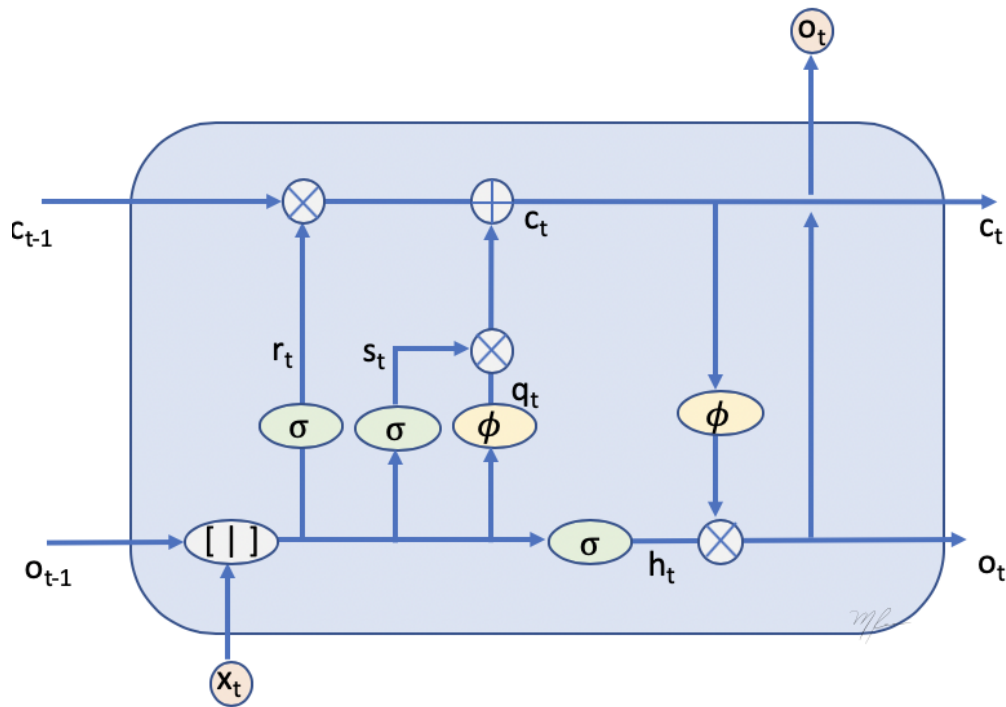


Figura 2.6: Estructura red neuronal LSTM.

- Puerta de reajuste: Incorpora las entradas actualizadas con los contenidos anteriores de la celda.

En comparación a las redes LSTM, que pueden operar con conjuntos de datos más grandes, por lo cual son más expresivas y proporcionar mejores resultados ante contextos complejos. Sin embargo, las redes GRU han demostrado que presentan un mejor rendimiento en conjuntos de datos más pequeños. Ya que son más simples de diseñar:

- Requieren de un tiempo de aprendizaje menor.
- Tienen menos hiperparámetros.
- Ejecución más eficiente.
- No requieren de puerta de salida.
- Permite modelar una red neuronal recurrente estándar, al establecer la puerta de reajuste a 1 y la puerta de actualización a 0.

Por lo que las redes neuronales recurrentes, son consideradas muy potentes, por esta razón se emplean para procesar secuencias de datos en aplicaciones como lo son: reconocimiento de voz, la traducción automática y la generación de texto.

2.9.2. Clasificación según el método de aprendizaje

Para una red neuronal; se derivan diferentes tipos de metodologías de aprendizaje; donde consiste en distintas etapas, aunque el contenido o la metodología utilizada varíe, las etapas suelen ser muy similares con diferencias entre sí.

Etapas 1. Recepción de la información.

Etapa 2. Transmisión de la información.

Etapa 3. Comprensión de información.

Etapa 4. Retención la información a largo plazo.

Y en estas mismas se pueden subdividir; estas mismas metodologías de aprendizaje como:

Aprendizaje supervisado

El proceso de aprendizaje es realizado por un entrenamiento controlado o supervisado [45], en el cual se determina la respuesta que debe de generar la entrada; esto se realiza a partir de que se expone el sistema a una entrada a la cual se conoce la respuesta de salida que se debe obtener, en el caso de que no sea correcta se modifican los pesos sinópticos, con el fin de que la salida sea lo más próxima a lo deseado.

El aprendizaje supervisado puede realizarse por medio de los siguientes aprendizajes:

- Aprendizaje por corrección de error

El cual ajusta los pesos de las conexiones de la red de acuerdo a la función de error cometido; el cual, es decir, que es la diferencia entre los valores esperados a los obtenidos. Donde algunos algoritmos que se utilizan son:

1. Perceptrón.
2. Mínimo error cuadrado (Least Mean Squared).
3. Programación hacia atrás (Backpropagation).

- Aprendizaje estocástico

Es por el cual se realizan cambios aleatorios sobre los pesos; donde se va calculando la predicción y a su vez va mejorando o empeorando con cada cambio hecho en los pesos; por el cual en su aprendizaje va quedándose con los cambios que mejor resultado den.

Aprendizaje no supervisado [46]

Este tipo de aprendizaje es caracterizado porque no se requiere de conocer la salida para el ajuste de los pesos sinópticos; el cual busca encontrar las características, regularidades, correlaciones o las categorías que se puedan establecer entre los datos presentados en las entradas. Lo principal que se toma en cuenta es en la interpretación de los datos de los cuales dependerán de la estructura del algoritmo de aprendizaje empleado.

Como en el aprendizaje supervisado, el no supervisado también se puede dividir en diferentes tipos de aprendizaje como lo son:

- Aprendizaje hebbiano

Este tipo de aprendizaje permite medir la familiaridad de las características de los datos entrada.

- Aprendizaje competitivo y comparativo

Este permite realizar la clasificación de los datos de entrada, donde consiste en ir añadiendo elementos a una clase; por el cual si un nuevo elemento se determina que es de la misma clase, los matiza; pero si son de forma contraria, crea una nueva clase con el elemento asociado a una serie de pesos propios.

- Aprendizaje por refuerzo

Es considerado el aprendizaje más lento de su estilo; dado que en este caso no se dispone de un conjunto complejo de datos exactos de salida, por el cual este indica si el dato es aceptable o no, donde el algoritmo ajusta los pesos que se basan en un mecanismo de probabilidades.

2.10. Comparación de redes neuronales

En la tabla 2.2, se presentan las características más relevantes de las redes neuronales anteriormente mencionadas, denotando sus ventajas y desventajas en cada caso.

Red	Ventajas	Desventajas
Conexión hacia adelante (red feedforward)	<ul style="list-style-type: none"> ■ Empleado regularmente para selección y la clasificación de objetos. ■ Permite el aislamiento de elementos del sistema para su análisis. 	<ul style="list-style-type: none"> ■ Requiere de un alto número de hiperparámetros. ■ Requiere de un tiempo de aprendizaje largo.
Recurrentes	<ul style="list-style-type: none"> ■ Empleada en predicción de series de tiempo, ya que presenta un buen desempeño. ■ Requiere del uso de la memoria interna. ■ El tiempo de aprendizaje es menor en comparación a la red feedforward. ■ Implementada en aplicaciones de identificación y/o control de sistemas. 	<ul style="list-style-type: none"> ■ Lo aprendido por la red, solo se puede manejar con memoria de largo ■ Igual que las redes feedforward, se requiere de un gran dimensionamiento. ■ Presenta problemas durante la transferencia de información en la fase de aprendizaje. ■ Algoritmos de aprendizaje más complejos en comparación a las redes feedforward, debido a la presencia de recurrencias.
Convolucionales	<ul style="list-style-type: none"> ■ Considera una versión mejorada de las redes feedforward. ■ La fase de aprendizaje se realiza más rápido en comparación de sus antecesoras. ■ Las dimensiones de la red son menores a comparación con la red feedforward. 	<ul style="list-style-type: none"> ■ Debido a la operación de convolucional, la fase de aprendizaje se complica, ya que requiere un aprendizaje más preciso. ■ Presenta mejores resultados cuando se emplea, varias estructuras de la misma red conectadas entre sí.

Híbridas

- Presenta mejores resultados cuando se emplea una sola red.
- Permite aprovechar las ventajas de la cada red, para así obtener mejores resultados al combinar las estructuras.
- Las desventajas de las estructuras de las redes empleadas, se pueden heredar total o parcialmente a la estructura diseñada.
- La fase de aprendizaje se complica proporcionalmente a la complejidad de la estructura de la red.

Multilayer perceptron (Perceptrón multicapa)

- Presenta una alta capacidad de aprendizaje.
- Tienen una estructura flexible, por lo cual se pueden diseñar, para realizar predicciones con datos de diversas naturalezas como: textos, imágenes, audio, señales de sensores, etc.
- En el caso de modelos no lineales son capaces de recrearlos o crearlos.
- Permite ejecutar la fase de aprendizaje online (en línea), a partir de mediciones que estén disponibles.
- Requiere una base de datos grande para realizar la fase de aprendizaje.
- Debido a su capacidad de aprendizaje, tiende a presentar sobre-aprendizaje con mucha facilidad.
- Presenta una función de pérdida en general no convexa con más de un mínimo local, por lo que los valores aleatorios inicialmente asignados a los pesos sinópticos de los enlaces entre neuronas pueden tender a modelos con diferente precisión.
- Requiere de un alto número de hiperparámetros, lo que dificulta su optimización.
- Son muy sensibles a la escala de los datos, por lo que estandarizarlo de tal forma que todas tengan un mismo valor medio e idéntica desviación estándar

2.11. Redes neuronales y sus aplicaciones

Las redes neuronales se han utilizado para aprender estrategias de solución, esto se debe a que pueden aprender y modelar las relaciones entre los datos de entrada y salida. Permitiendo su uso en las siguientes tareas [47]:

- Procesamiento de imágenes, para el diagnóstico médico mediante la clasificación de imágenes médicas.
- Reconocimiento de patrones, empleado en marketing mediante el filtrado de redes sociales y el análisis de datos de comportamiento.
- Predicciones, empleado en el sector financiero mediante el procesamiento de datos históricos de instrumentos financieros.
- Filtrado de señales.
- Procesos en donde se requiera un control de calidad o una optimización.
- Visión artificial, con las redes neuronales, la computadora puede distinguir y/o reconocer imágenes.
 - Reconocimiento visual de señales de tráfico y a otros usuarios del camino en los vehículos autónomos.
 - Moderación de contenido para eliminar contenido inseguro o inapropiado en los archivos de imágenes y videos.
 - Identificación de rostros y reconocimiento de rasgos distintivo como ojos abiertos, gafas y vello facial, mediante el uso de reconocimiento facial.
 - Identificación de logotipos de marcas, ropa, equipos de seguridad en imágenes.
- Reconocimiento de voz, mediante el análisis del habla humana a pesar de los diferentes patrones de habla, el tono, el idioma y el acento.
 - clasificación de llamadas de forma automática en centros de llamadas.
 - Obtención de documentos de textos a partir de conversaciones en tiempo real.
 - Generación de subtítulos en videos y/o grabaciones de reuniones.
- Procesamiento de lenguaje natural, procesamiento de los datos y los documentos de texto para obtención de información.
 - Generación de chatbots y agentes virtuales automáticos.
 - Organización y clasificación de datos escritos.
 - Análisis de documentos como lo son: formato largo, emails y formularios.
 - Generación de resúmenes a partir de documentos.
 - Clasificación de frases de acuerdo a palabras clave.

2.11.1. Ventajas

Las redes neuronales permiten la toma de decisiones inteligentes en procesos computacionales con limitada participación humana, esto debido a que se encuentran diseñadas para aprender y modelar relaciones complejas y no lineales entre datos de entrada y salida, permitiendo denotar las siguientes ventajas [48].

- La red neuronal presenta robustez que permite reconocer patrones con ruido, distorsionados o incompletos en datos de entrada y salida.
- Ante la destrucción parcial de la red, se puede retener información que ha sido procesada antes de presentarse la degradación de la estructura.
- La complejidad del problema y los criterios de diseño de la red son independientes entre sí, debido a que se puede utilizar una red neuronal con el mismo número de neuronas y/o capas ocultas para problemas de diferentes complejidades y lo que determinara si se requiere modificar sus criterios de diseño serán los índices de desempeño obtenidos.
- Los criterios de diseño son de rápido ajuste y simulación que permite obtener gráficas en las que se puede observar desempeño que se tiene en cada intención de entrenamiento o la actualización de los pesos sinópticos de cada neurona.
- Presenta un aprendizaje adaptativo, el cual consiste en aprender a realizar tareas basadas a partir de una fase de aprendizaje o una experiencia inicial.
- Presenta un proceso de auto-organización mediante el cual crea su propia organización y/o representación de información que recibe durante la etapa de aprendizaje.
- Responden apropiadamente ante la generalización, esto significa que presenta la habilidad de inducir características generales a partir de los datos particulares de aprendizaje, permitiendo reconocer patrones a los que no han sido presentados con anterioridad.
- Permite operación en tiempo real.
- Las redes neuronales presentan una inserción dentro de la tecnología existente, debido a que mejoran su funcionalidad en cierta tarea.

2.11.2. Desventajas

Así mismo, las redes neuronales, ante la falta de reglas definitorias durante el procedimiento de realización de una red, se presentan algunos problemas que hacen notar las desventajas [49] de las mismas, como lo son:

- Presenta un comportamiento de un sistema de black box (caja negra).
- Los criterios de diseño son arbitrarios, entre estos se encuentran: el dimensionamiento, inicialización de pesos sinópticos, número de bias.
- Soluciones diferentes ante la división aleatoria de datos para aprendizaje, validación y test, que pueden provocar sobreaprendizaje o pérdida de generalización.
- La complejidad de aprendizaje está relacionada con la tarea a realizar, debido a que entre más necesite aprender una red, más complicado será entrenarla.

- El tiempo de aprendizaje depende de dos factores: la cantidad de patrones a identificar o clasificar y la flexibilidad o capacidad de adaptación de la red neuronal para reconocer patrones que sean sumamente parecidos, esto determinará el tipo requerido para que la red converja.
- Cuando se requiere una mayor flexibilidad para la red neuronal, se necesita una cantidad mayor de datos para su aprendizaje debido a que se necesita más información para enseñarle y así realice de forma adecuada la identificación.
- La red por sí sola proporciona una salida, que no puede ser interpretado por sí misma, por lo que requiere de la intervención del programador y de la aplicación en sí para interpretar lo que se ha aprendido y así encontrar un significado a la salida proporcionada.

Por lo que, la red neuronal más común para este tipo de estudio es la implementación de red neuronal multicapa entrenada por backpropagation, también conocida como «red neuronal backprop». La red neuronal backprop con una función de activación sigmoide en cascada donde las salidas de una capa, multiplicadas por los pesos, sumadas y luego desplazadas por un sesgo, se utilizan como las entradas de la capa siguiente. Una vez elegida la arquitectura de la red, la función representada por la red neuronal se codifica mediante los pesos y los sesgos.

De modo que datos se ejecutan repetidamente a través del aprendizaje, hasta llegar a un error aceptable entre en los datos y la red neuronal. Lo que nos lleva a que las ventajas, limitaciones y desventajas de esta metodología pueden ser englobadas juntas debido a que depende de la estructura del caso de estudio:

- El peso computacional asociado al aprendizaje de la red neuronal está directamente relacionada con el tamaño de la estructura, por lo cual al tener mayor número de pisos la estructura, el aprendizaje de la red es más mayor.
- Se requiere utilizar el algoritmo de optimización adecuado al sistema con tal de obtener un mejor procesamiento de datos.
- La definición de la cantidad de capas ocultas y el número de neuronas por capa se realizó a partir de diferentes simulaciones.
- Selección adecuada de datos de aprendizaje para la red neuronal.
- La red neuronal puede presentar sobre-aprendizaje, en este sentido la red presenta un aprendizaje a tal grado que solo puede identificar un sistema siempre y cuando las condiciones de aprendizaje y prueba sean iguales o similares lo más posible.

Capítulo 3

Modelo matemático de las estructuras de tipo edificio

3.1. Conceptos fundamentales

Antes de presentar el modelo matemático recordaremos algunos conceptos para entender los resultados obtenidos. Debido a que las estructuras civiles experimentan vibraciones ante la presencia de actividad sísmica, se obtiene una base de datos para determinar la frecuencia natural y la relación con el amortiguamiento de la misma.

3.1.1. Vibración

Una vibración mecánica [50] se describe como el movimiento periódico repetitivo de un objeto alrededor de una posición de equilibrio, sin que se genere desplazamiento.

- Vibración libre: Se produce a partir de un impulso inicial con condiciones iniciales en posición y velocidad.
- Vibración forzada: Se produce partir de una fuerza o perturbación externa al sistema, aplicada de forma continua en cada instante de tiempo.

Las vibraciones que se presentan a partir de terremoto en las estructuras civiles suelen provocar daños a la construcción, como lo son:

- Destrucción del techo.
- Vidrios que se rompen.
- Fisuras en pisos y columnas.

El criterio empleado para la identificación de daños generados por la vibración está enfocado en la deformación de materiales. Por lo que la resistencia de materiales es el punto admisible para localizar la deformación en las estructuras civiles estudiadas, y se pueda concluir el límite admisible de la vibración que soporta el cuerpo de estudio antes de presentar algún daño.

3.1.2. Movimiento armónico

Es un movimiento vibratorio donde la posición ($x(t)$), velocidad ($v(t)$) y aceleración ($a(t)$) suelen ser descritas mediante funciones senoidales o cosenoidales, generalmente en una sola dirección.

El movimiento armónico más sencillo es el movimiento armónico simple (ver figura 3.1), donde la ecuación de movimiento es, dada por una función:

$$\begin{aligned} x(t) &= A\sin(\omega t + \phi) \\ x(t) &= A\cos(\omega t + \phi) \end{aligned} \tag{3.1}$$

donde, A , ω , t y ϕ , representan la amplitud de la onda, frecuencia excitante, instante en un tiempo y ángulo de fase, respectivamente.

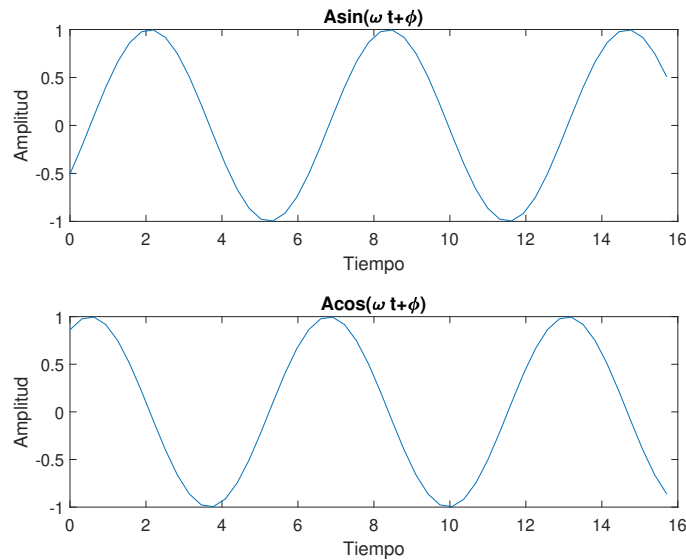


Figura 3.1: Movimiento armónico.

3.1.3. Frecuencia

La frecuencia de un objeto en vibración se determina con el número de ciclos completos, generados en un segundo, siendo su unidad de medida Hertz (Hz) [51].

$$\begin{aligned} T &= \frac{1}{f} \longrightarrow T = n \cdot d_t \\ F &= \frac{\omega_n}{2\pi} \end{aligned} \tag{3.2}$$

donde, T , f , n , d_t , ω_n , π , representan el periodo, la frecuencia, el número de ciclos completos, el tiempo de muestreo, la frecuencia natural y la relación entre la longitud de una circunferencia con su diámetro en geometría euclidiana respectivamente.

- Frecuencia natural [52]: Es la frecuencia a la que sigue vibrando un sistema mecánico, después que se retira la señal de excitación.

$$\omega_n = 2\pi f \tag{3.3}$$

- Frecuencia de excitación externa: En el caso de las estructuras civiles, las fuentes de vibraciones externas son generalmente debidas a: circulación vehicular en calles o rutas y ferrocarriles, operación de maquinaria pesada relacionada con la construcción, vientos fuertes, así como los movimientos telúricos o comúnmente conocidos como actividad sísmica.

$$x(t) = \frac{1}{\omega_n - \omega_e} \quad (3.4)$$

- Frecuencia de resonancia: La Resonancia es un fenómeno que amplifica una vibración. Se produce cuando una vibración externa se transmite a un objeto cuya frecuencia natural es igual o muy cercana a la de la fuente.

$$\omega_e \approx \omega_n \quad (3.5)$$

3.1.4. Amortiguamiento

El amortiguamiento en las estructuras civiles [53] es considerado un efecto disipativo, el cual se representa con la pérdida de energía que genera un sistema mecánico en movimiento. Siendo un fenómeno asociado con la fricción y las micro-fracturas internas.

Los valores de amortiguamiento en estos modelos son aproximados y se basan en mediciones experimentales en las estructuras civiles. Por lo que todas las estructuras poseen un factor de amortiguamiento propio, pero hay dispositivos que incrementan el amortiguamiento disipando energía cuando la estructura se mueve.

Las tres fuentes típicas de amortiguamiento [54] son:

- La fricción seca o rozamiento.
- La fricción viscosa.
- El amortiguamiento histerético asociado a la deformación de los materiales.

3.2. Instrumentos de medición de vibraciones

La medición de vibraciones es de gran interés en la ingeniería estructural, ya que se requiere el estudio de las oscilaciones mecánicas que se presentan en las estructuras civiles; por consecuencia la medición del movimiento del suelo durante un terremoto proporciona datos básicos con la finalidad de registrar los movimientos resultantes de una estructura brindando la información sobre el cómo se comporta durante el movimiento telúrico, para realizar estudios, análisis o conclusión del sistema bajo pruebas.

El proceso de medición se ilustra en la figura 3.2 [55], donde se muestra el procedimiento para la adquisición de la medición, el cual se encuentra conformado por:

- Transductores de vibraciones: Se representa como un sistema de masa-resorte-amortiguador montado dentro de un marco rígido que se fija a la superficie cuyo movimiento se va a medir. Funcionando a partir del movimiento del punto de apoyo, la masa del transductor se mueve en relación con el marco y el desplazamiento relativo se registra en la señal obtenida, trasforma las vibraciones en señales eléctricas.
- Un sistema de acondicionamiento de señal: Recolecta las diferentes señales para amplificarlas a niveles de tensión adecuados para el sistema de adquisición de datos.

- Tarjeta de adquisición de datos: Realiza la digitalización de la señal, en la cual se realiza un muestreo individual de la señal analógica obtenida, en el acondicionamiento de señal para así ingresarla a una computadora donde se realizan diversos tipos de procesamiento, obtener toda la información necesaria para análisis y/o control de vibraciones.

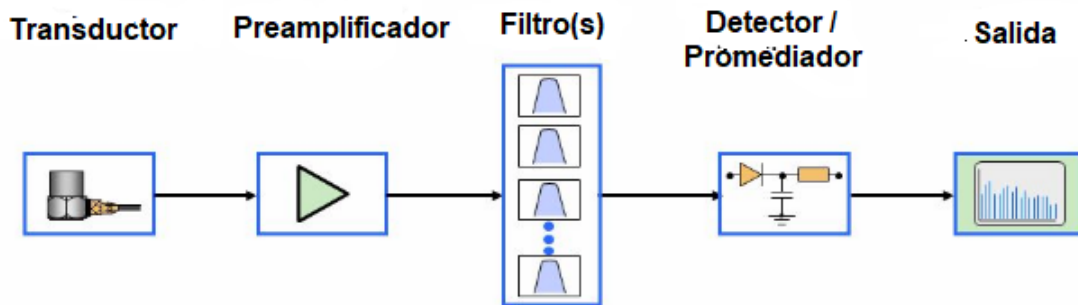


Figura 3.2: Proceso de Medición de vibraciones.

Los datos obtenidos pueden ser usados para minimizar o eliminar el ruido de medición generado en la obtención de las mediciones de vibración.

3.2.1. Medición de desplazamiento

La utilización de mediciones de desplazamiento [56] como factor para determinar el deterioro o daño sufrido en estructuras civiles a partir de la presencia de actividad sísmica, son poco empleadas debido a la complejidad para obtener instrumentos de medición que proporcionen las correctas mediciones, así como de la idealización en los escenarios para su utilización.

Debido a esto, la medición de desplazamiento requiere tener disponible el desplazamiento relativo ($u(t)$) y el desplazamiento del suelo o soporte ($u_g(t)$) en caso de ser un modelo experimental. Estas mediciones se basan en hacer que el resorte del transductor de vibración que se coloque en el soporte, sea menos flexible o que la masa del transductor sea mayor para poder obtener una lectura.

Pero si se emplean ambas consideraciones al momento de presentarse actividad sísmica, el soporte oscilará y la mesa a la cual se encuentre sujeta no se moverá. Siendo así un instrumento de medición difícil de manejar, debido a las características que debe cumplir el transductor de vibración; ya que se debe anticipar el desplazamiento máximo del soporte (u_{go}), el cual puede oscilar entre 12 y 36 pulgadas dependiendo de la intensidad del movimiento telúrico. Por lo que el movimiento armónico ante actividad sísmica queda descrito como:

$$u_g(t) = u_{go} \sin \omega t \tag{3.6}$$

Para frecuencias de excitación superiores a la frecuencia natural; que presenten un factor de modificación en la fuerza (R) cerca de la unidad y el ángulo de fase (ϕ) sea aproximadamente de 180° , y la ecuación se reescribe como:

$$u(t) = -u_{go} \sin \omega t \tag{3.7}$$

Por lo cual, el desplazamiento registrado es igual al desplazamiento del soporte con signo negativo. Pero si la frecuencia de resonancia es muy pequeña, la atenuación del instrumento no es un parámetro crítico, por lo cual tendrá poco efecto en el movimiento registrado.

3.2.2. Medición de aceleración

El movimiento a medir suele variar aleatoriamente con el tiempo y puede contener muchos componentes armónicos que cubren una amplia gama de frecuencias. Sin embargo, es muy importante considerar primero el grado de movimiento armónico, que generará el desplazamiento de la masa estructural con respecto a la estructura metálica o marco móvil en prototipos se puede reescribir como:

$$u(t) = - \left[\left(\frac{1}{\omega_n^2} \right) R_d \right] \ddot{u}_g \left(\frac{\phi}{\omega} \right) \quad (3.8)$$

donde, R_d y \ddot{u}_g son dos términos nuevos que representan el factor de respuesta de deformación y la aceleración del suelo respectivamente.

En la ecuación anterior, el desplazamiento se registra por medio de la aceleración con un factor de corrección dado por $-R_d/\omega^2$ y el registro por el desfase temporal es ϕ/ω . Y esto depende de la frecuencia ω , pero ω_n^2 es una constante en el instrumento independiente del movimiento del soporte.

El diseño se basa en hacer que el factor de modificación en la fuerza (R), y ϕ/ω sean independientes de la frecuencia de excitación. Debido a que cada componente armónico de la aceleración [57] se registra con el mismo coeficiente relación y con un retardo de tiempo.

Aunque el movimiento registrado consta de muchos componentes armónicos, el desplazamiento registrado tiene la misma forma que el movimiento de referencia en el tiempo.

Los equipos disponibles actualmente para medir la aceleración del suelo debido a terremotos, presentan las siguientes características:

- El rango de frecuencia es de 0 a 100 Hz y el error es insignificante.
- Se registra el desplazamiento del suelo y no la aceleración.
- El rango de movimiento del suelo es desde 0.001 hasta 2 g , donde g es la aceleración gravitacional, que siente un cuerpo debido a la fuerza de la gravedad

3.2.3. Información requerida para la medición

En la medición de prueba de vibración se debe tener en cuenta diversos factores los cuales determinaran la metodología a utilizar, donde se consideran:

- Selección de instrumentación: se considera que en la mayoría de los casos, las mediciones no se pueden tomar debido a restricciones del terreno que impiden el acceso a la estructura. Lo cual se complica debido a que los dispositivos requieren estar en contacto con la superficie que se está midiendo, no como en el caso de los acelerómetros, pero siendo esta la única medición disponible.
- Características de la fuente de excitación:
 - Duración de la excitación.
 - Variación de la amplitud de la excitación: Estable, cíclica, aleatoria o aperiódica.
- Información primaria concerniente a la estructura específica que se va a estudiar como son:
 - Características estructurales: Altura, extensión, materiales y fundación.

- Parámetros dinámicos teóricos: Formas de modales, frecuencias naturales y amortiguamiento.
 - Características del terreno: Propiedades geotécnicas del suelo, excavaciones, estabilidad.
- Tipo de análisis para los resultados.

3.3. Modelo matemático

El modelado matemático es hoy un medio imprescindible para la resolución de una gran variedad de problemas (científicos, económicos, biológicos) y en particular para el estudio de los llamados sistemas dinámicos. Su estudio está motivado por los avances en procesadores de cálculo, nuevas teorías matemáticas y el análisis de sistemas [58].

Desde el punto de vista matemático, los modelos se definen como descripciones lógicas sobre hechos y/o fenómenos del mundo real, que van desde el crecimiento de una población hasta fenómenos físicos como la velocidad, como el voltaje o la temperatura.

La finalidad del modelado matemático es entender ampliamente el fenómeno y así predecir o tener una idea de su comportamiento futuro bajo determinadas condiciones, los puntos del proceso en general para elaborar un modelado matemático que involucren los siguientes pasos [59]:

1. Seleccionar un problema del mundo real: El modelado matemático es en cierta medida multidisciplinario, ya que bien puede aplicarse tanto a la ingeniería en sus diferentes ramos, como a la economía, el medio ambiente, astronomía, la agricultura, etc.
2. Formular un modelo matemático acerca del problema: Mediante la identificación de las variables del problema como la aplicación de las leyes universales (físicas, químicas, matemáticas) estableciendo hipótesis que permitan tratar dicha formulación de manera matemática.
3. Aplicar los conocimientos matemáticos de abstracción: De tal forma que permitan obtener conclusiones matemáticas concretas (modelos matemáticos reducidos).
4. Simular y comparar los resultados obtenidos con los datos reales: Esto para validar el modelo matemático resultante. Si los datos difieren considerablemente, se reinicia el proceso.

Para solucionar el problema de modelado se entiende la comprensión o entendimiento de un sistema determinado que permita predecir su comportamiento y posiblemente obtener información sobre cómo identificar y/o controlar el sistema real. En general, los modelos de sistemas se pueden obtener de 2 formas:

- Analíticamente.
- Experimentalmente.

Algunos de los problemas principales de obtener modelos mediante experimentos son los siguientes:

- Podría resultar muy caro.

- Podría resultar muy peligroso.
- El sistema requerido para el experimento podría no existir.

Por otro lado, coma cuando el modelo se obtiene a través del análisis (matemático) del sistema, entonces se recurre a la simulación. La simulación es un experimento desarrollado sobre un modelo virtual, con el objetivo de estudiar su comportamiento bajo diferentes condiciones de operación. Algunas de las razones principales para utilizar la simulación son:

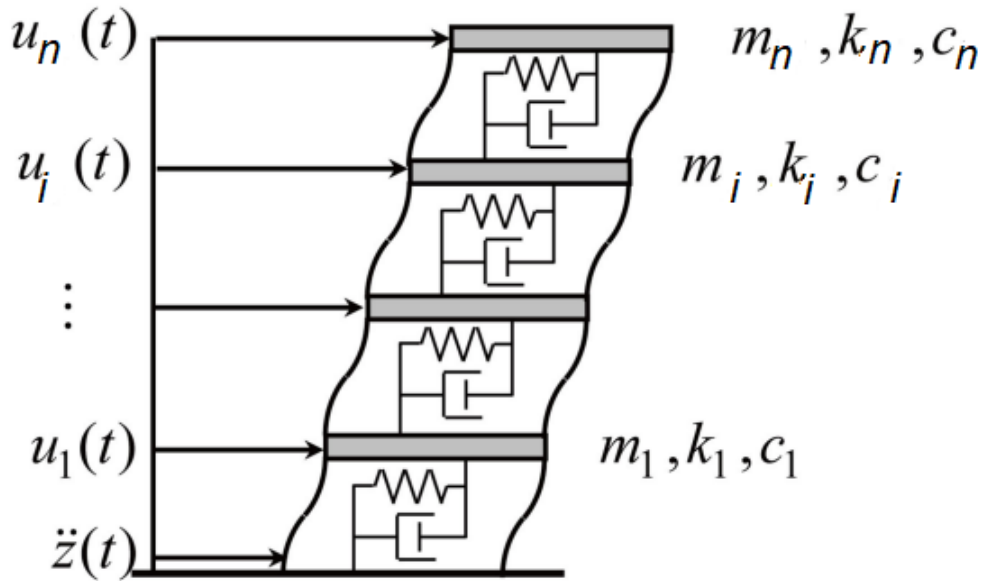
- Supresión de efectos secundarios.
- Se evita el problema de hacer experimentos muy caros o muy peligrosos.
- La escala de tiempo no es compatible con el experimento.
- Las variables a medir podrías estar e inaccesibles.
- Fácil manipulación.
- Supresión de perturbaciones.

Sin embargo, el abuso en la simulación también trae consigo sus consecuencias. Algunos de los problemas que se tienen al utilizar la simulación son:

- Forzar la realidad a las restricciones del modelo.
- Olvidar que el modelo tiene un nivel de precisión limitado.

Ahora bien, empleando un formalismo del modelo matemático [60] consiste en la representación de un sistema o proceso por medio del lenguaje matemático, de tal modo que describe y predice el comportamiento de las características más sobresalientes y de mayor interés. Donde los elementos esencialmente son las variables y las funciones relacionadas mediante expresiones matemática.

Por consiguiente, si aplicamos lo anterior para obtener un modelo matemático que describa el comportamiento de una estructura civil, como un sistema lineal de múltiples grados de libertad, como se presenta en la figura 3.3 [61], la cual representa una edificación sometida a una excitación externa de tipo sísmica.


 Figura 3.3: Edificio de n – pisos.

Para obtener el modelo matemático de una estructura de tipo inmueble, esta se considera como un conjunto de n masas que se encuentran interconectadas por medio de resortes, así como de amortiguadores viscosos lineales. Donde se tienen las siguientes consideraciones:

- Los cimientos de la estructura se encuentra enterrados bajo suelo firme, para así soportar el peso de la edificación.
- Los diafragmas o pisos de una edificación, determinan el comportamiento de las cargas gravitacionales y laterales que presenta una estructura. Por lo cual tiene dos funciones:
 1. Transferir las cargas horizontales o cargas sísmicas a los elementos verticales.
 2. Transferir cargas verticales a los pisos inferiores, a través de las columnas que son flexibles para deformación.

Por lo cual se emplea la formulación de Euler-Lagrange que incluye funciones de energía disipativa, requeridas para las estructuras tipo edificio, lo que nos lleva a la ecuación (3.9).

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}_i} \right) - \frac{\partial L}{\partial x_i} + \frac{\partial R}{\partial \dot{x}_i} = F_e \quad (3.9)$$

Donde:

- F_e representa las fuerzas externas ejercidas en la estructura tipo edificio por un sismo.
- Las coordenadas generalizadas son x_i $i = 1, 2, \dots, n$.
- R es una función de disipación de Rayleigh donde en la que se representa la fuerza absorbida por el amortiguador por la estructura tipo edificio [62].
- L es la función lineal conocida como lagrangiano, la cual se define en la ecuación (3.10) que involucra T y U las cuales representan la energía cinética y la energía potencial, respectivamente.

$$L = T - U \quad (3.10)$$

Las funciones T , U , y R , son definidas en las ecuaciones siguientes (3.11), (3.12) y (3.13), en función de las variables m_i , c_i y k_i con $i = 1, 2, \dots, n$, son la i -ésima masa, rigidez y amortiguamiento de entre los pisos de la estructura, siendo estas las características físicas.

$$T = \frac{1}{2}(m_i x_i^2) \quad (3.11)$$

$$U = \frac{1}{2}k_i(x_i - x_{i-1})^2 \quad (3.12)$$

$$R = \frac{1}{2}c_i(\dot{x}_i - \dot{x}_{i-1})^2 \quad (3.13)$$

Tal que se inicia desde $i = 1$ hasta n :

$$T = \sum_{i=1}^n \frac{1}{2}(m_i x_i^2) = \frac{m_1 \dot{x}_1^2}{2} + \frac{m_2 \dot{x}_2^2}{2} + \frac{m_3 \dot{x}_3^2}{2} + \dots + \frac{m_n \dot{x}_n^2}{2} \quad (3.14)$$

$$U = \sum_{i=1}^n \frac{1}{2}k_i(x_i - x_{i-1})^2 = -\frac{k_1}{2}[x_1]^2 - \frac{k_2}{2}[x_2 - x_1]^2 - \frac{k_3}{2}[x_3 - x_2]^2 - \dots - \frac{k_n}{2}[x_n - x_{n-1}]^2 \quad (3.15)$$

$$R = \sum_{i=1}^n \frac{1}{2}c_i(\dot{x}_i - \dot{x}_{i-1})^2 = \frac{c_1}{2}[\dot{x}_1]^2 + \frac{c_2}{2}[\dot{x}_2 - \dot{x}_1]^2 + \frac{c_3}{2}[\dot{x}_3 - \dot{x}_2]^2 + \dots + \frac{c_n}{2}[\dot{x}_n - \dot{x}_{n-1}]^2 \quad (3.16)$$

El siguiente paso es definir las ecuaciones de movimiento, a partir de las derivadas parciales de la ecuación (3.9), con respecto a la primera coordenada generalizada x_i , como se expresan a continuación para el caso x_1 se obtienen las ecuaciones (3.17), (3.18) y (3.19).

$$\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial \dot{x}_1} = m_1 \dot{x}_1 \quad (3.17)$$

$$\frac{\partial L}{\partial q_i} = \frac{\partial L}{\partial x_1} = -k_1(x_1) + k_2(x_2 - x_1) \quad (3.18)$$

$$\frac{\partial R}{\partial \dot{q}_i} = \frac{\partial R}{\partial \dot{x}_1} = c_1(\dot{x}_1) + c_2(\dot{x}_1 - \dot{x}_2) \quad (3.19)$$

Se realiza el mismo procedimiento, para la coordenada x_2 :

$$\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial \dot{x}_2} = m_2 \dot{x}_2 \quad (3.20)$$

$$\frac{\partial L}{\partial q_i} = \frac{\partial L}{\partial x_2} = -k_2(x_2 - x_1) + k_3(x_3 - x_2) \quad (3.21)$$

$$\frac{\partial R}{\partial \dot{q}_i} = \frac{\partial R}{\partial \dot{x}_2} = c_2(\dot{x}_2 - \dot{x}_1) + c_3(\dot{x}_2 - \dot{x}_3) \quad (3.22)$$

Se repite este procedimiento para obtener las derivadas de cada piso x_i de la estructura hasta llegar al último piso x_n .

$$\frac{\partial L}{\partial \dot{q}_i} = \frac{\partial L}{\partial \dot{x}_n} = m_n \dot{x}_n \quad (3.23)$$

$$\frac{\partial L}{\partial q_i} = \frac{\partial L}{\partial x_n} = -k_n(x_n - x_{n-1}) \quad (3.24)$$

$$\frac{\partial R}{\partial \dot{q}_i} = \frac{\partial R}{\partial \dot{x}_n} = c_n(\dot{x}_n - \dot{x}_{n-1}) \quad (3.25)$$

Al resolver las derivadas con respecto al tiempo se obtiene un conjunto de ecuaciones diferenciales ordinarias (3.17)-(3.25), que se sustituyen en la ecuación dinámica de Euler-Lagrange (3.9).

$$m_1\ddot{x}_1 + c_1\dot{x}_1 + k_1x_1 + k_2(x_1 - x_2) + c_2(\dot{x}_1 - \dot{x}_2) = F_{e1} \quad (3.26)$$

$$m_2\ddot{x}_2 - c_2\dot{x}_1 + (c_2 + c_3)\dot{x}_2 - c_3\dot{x}_3 - k_2x_1 + (k_2 + k_3)x_2 - k_3x_3 = F_{e2} \quad (3.27)$$

$$m_3\ddot{x}_3 - c_3\dot{x}_2 + (c_3 + c_4)\dot{x}_3 - c_4\dot{x}_4 - k_3x_2 + (k_3 + k_4)x_3 - k_4x_4 = F_{e3} \quad (3.28)$$

$$\begin{aligned} & \vdots \\ & = \vdots \\ m_n\ddot{x}_n + k_n(x_n - x_{n-1}) + c_n(\dot{x}_n - \dot{x}_{n-1}) &= F_{en} \end{aligned} \quad (3.29)$$

Donde:

- F_{ei} , representa la actividad sísmica que afecta a cada piso de la estructura y su comportamiento se describe por la ecuación (3.30).

$$F_{ei} = m_i\ddot{u}_g(t) \quad (3.30)$$

- \ddot{u}_g , es la aceleración del suelo en el que se encuentre fijada la estructura.

El sistema puede ser reescrito en forma matricial con M_s , C_s y K_s , como se presenta en la ecuación (3.31), a partir de las ecuaciones (3.26) - (3.29).

$$M_s\ddot{x}(t) + C_s\dot{x}(t) + K_sx(t) = -M_s1\ddot{x}_g \quad (3.31)$$

$$1 = [1, 1, 1, \dots, 1] \in \mathbb{R}^{n \times 1}$$

Donde, 1 es el vector que distribuye la señal escalar \ddot{x}_g a cada piso de la estructura. Y las matrices de masa, amortiguamientos y rigidez, se encuentran definidas por las ecuaciones (3.32)-(3.34), en las cuales m_i , c_i y k_i con $i = 1, 2, \dots, n$, son la i -ésima masa, rigidez y amortiguamiento de entre los pisos de la estructura, siendo estas las características físicas.

$$M_s = \text{diag}[m_1, m_2, m_3, \dots, m_n] > 0 \in \mathbb{R}^{n \times n} \quad (3.32)$$

$$C_s = \begin{bmatrix} c_1 + c_2 & -c_2 & \cdots & 0 & 0 \\ -c_2 & c_2 + c_3 & \cdots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \cdots & c_{n-1} + c_n & c_n \\ 0 & 0 & \cdots & -c_n & c_n \end{bmatrix} > 0 \in \mathbb{R}^{n \times n} \quad (3.33)$$

$$K_s = \begin{bmatrix} k_1 + k_2 & -k_2 & \cdots & 0 & 0 \\ -k_2 & k_2 + k_3 & \cdots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \cdots & k_{n-1} + k_n & k_n \\ 0 & 0 & \cdots & -k_n & k_n \end{bmatrix} \geq 0 \in \mathbb{R}^{n \times n} \quad (3.34)$$

Así mismo, los vectores de desplazamiento, velocidad y aceleración, son mediciones relativas con respecto a la posición inicial y están representados por $x(t)$, $\dot{x}(t)$ y $\ddot{x}(t)$, respectivamente.

$$\begin{aligned} x(t) &= [x_1(t), x_2(t), x_3(t), \dots, x_n(t)]^T \in \mathbb{R}^{n \times 1} \\ \dot{x}(t) &= [\dot{x}_1(t), \dot{x}_2(t), \dot{x}_3(t), \dots, \dot{x}_n(t)]^T \in \mathbb{R}^{n \times 1} \\ \ddot{x}(t) &= [\ddot{x}_1(t), \ddot{x}_2(t), \ddot{x}_3(t), \dots, \ddot{x}_n(t)]^T \in \mathbb{R}^{n \times 1} \\ 1 &= [1, 1, 1, \dots, 1] \in \mathbb{R}^{n \times 1} \end{aligned} \quad (3.35)$$

Definiendo la estructura tipo edificio en variables de estado se parte de la ecuación (3.36), lo que permite reescribir el modelo presentado en la ecuación (3.9), a la su representación en variables de estado como se presenta en la ecuación (3.37).

$$z(t) = \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} = \begin{bmatrix} x(t) \\ \dot{x}(t) \end{bmatrix} \quad (3.36)$$

$$\begin{bmatrix} \dot{z}_1(t) \\ \dot{z}_2(t) \end{bmatrix} = \underbrace{\begin{bmatrix} 0_{n \times n} & I_{n \times n} \\ -M_s^{-1}K_s & -M_s^{-1}D_s \end{bmatrix}}_A \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} - \underbrace{\begin{bmatrix} 0_{n \times n} \\ 1 \end{bmatrix}}_B \quad (3.37)$$

$$y(t) = \underbrace{\begin{bmatrix} I_{n \times n} & 0_{n \times n} \end{bmatrix}}_C z(t)$$

A partir del modelo en variables de estado de la estructura tipo edificio presentado en la ecuación (3.37), se obtiene que $y(t)$, representa las salidas del sistema en términos de desplazamiento que se presenta en cada piso. Esta representación del sistema permite:

- Realizar simulaciones a lazo abierto para la recolección de datos de comportamiento ante la presencia de una perturbación sísmica.
- La matriz A de la ecuación (3.37), proporciona los valores propios del sistema.
- Los polos del sistema se pueden expresar en términos de las frecuencias naturales y los factores de amortiguamiento, como se presenta en la ecuación (3.38). Donde $j \in \{1, 2, \dots, n\}$, de forma que para cada par de polos se pueda obtener la frecuencia natural y su respectivo factor de amortiguamiento.

$$s_{(2j-1, 2j)} = -\omega_{n(j)}\xi_{(j)} \pm i\omega_{n(j)}\sqrt{1 - \xi_{(j)}^2} \quad (3.38)$$

- En este tipo de sistemas que describen el comportamiento de una estructura de tipo edificio, los polos serán complejos conjugados [63].

3.3.1. Características que pueden ser modificadas

Las características que se requieren considerar y que pueden ser modificadas en el modelo son [64]:

- Las estructuras civiles reales son sistemas de múltiples grados de libertad, por lo cual suelen poseer el mismo número de formas modales que grados de libertad, donde cada una tiene su propia frecuencia a la que vibra ante presencia de actividad sísmica.
- La frecuencia fundamental es la forma modal más baja a la que vibra la estructura.
- El amortiguamiento de la estructura civil depende del tipo de construcción [65].
- El amortiguamiento suele aumentar con el incremento de los desplazamientos que presente la estructura civil.
- Los daños causados por vibración pueden ser mitigados al modificarse el periodo de vibración en la estructura, esto puede ser por medio de:
 - Aumento de masa [66]

Ejemplo 3. *Aumento de amortiguamiento por medio de masa sintonizada.*

Esto consiste en suspender un cuerpo esférico que se encontrará sujeto mediante cables de acero a la parte superior de la estructura, asemejando a péndulo y conectando la estructura con la esfera por medio de amortiguadores de fluido viscoso. El ejemplo más claro sería la Torre Taipei 101, en Taiwan como se presenta en la figura 3.4 [67], donde el péndulo se ajusta con precisión a la frecuencia natural de vibración de la estructura.

Esto es posible, ya que se emplea un acoplamiento entre la frecuencia natural de la estructura y un oscilador simple para contrarrestar la frecuencia de resonancia y minimizará la respuesta dinámica de la estructura.



Figura 3.4: Aumento de amortiguamiento por medio de masa sintonizada, en Torre Taipei 101.

- Amortiguamiento artificial [68].

Ejemplo 4. *Amortiguadores viscosos.*

Un amortiguador viscoso o de fluidos, se presenta en la figura 3.5 [69], el cual es un pistón que transmite la energía por medio del moviendo de fluido en el regulador, este movimiento del fluido absorbe la energía cinética y la convierte en calor. Por lo que, durante un terremoto, la fuerza generada por la aceleración se transmite al amortiguador, donde se regula el paso de fluido comprimido a través de pequeños hoyos. Así la energía sísmica es disipada tan rápido como el líquido fluye a través de los agujeros.

Esto genera que la estructura presenten menos movimientos horizontales, así como daños causados por actividad sísmica.

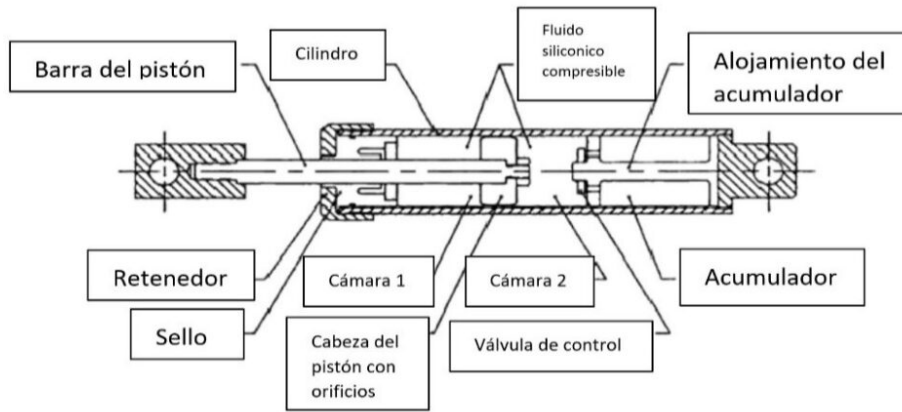


Figura 3.5: Funcionamiento de un amortiguador viscoso.

- Una estructura civil puede llegar a vibrar en frecuencias en las que un ser humano las perciba (3 Hz – 8 Hz).
- Una carga dinámica puede hacer resonar la estructura.

3.4. Modelo de histéresis de Bouc-Wen

El modelo de histéresis de Bouc-Wen [70] es el modelo matemático más utilizado en ingeniería estructural, debido a que describe el comportamiento no lineal de los materiales cuando estos se encuentran sometidos bajo cargas cíclicas, las cuales son generadas cuando las estructuras están sujetas a diferentes fuerzas mecánicas, deformaciones, tensiones y estrés. Este modelo se basa en la teoría de histéresis desarrollada por Bouc y Wen en la década de 1960, y desde entonces ha sido ampliamente utilizado en el análisis de estructuras sometidas a cargas sísmicas.

Se reescribe el sistema (3.31) como:

$$M_s \ddot{x}(t) + C_s \dot{x}(t) + K_s x(t) + \rho(x, z) = -M_s 1 \ddot{x}_g \quad (3.39)$$

$$\rho(x, z) = [\rho_1(x_1, z_1) \rho_2(x_2, z_2) \dots \rho_i(x_i, z_i)]^T$$

El modelo de histéresis de Bouc-Wen se basa en una serie de ecuaciones, las cuales describen la relación que existe entre la fuerza, el desplazamiento y la velocidad. Estas ecuaciones incluyen parámetros que representan la rigidez, el amortiguamiento, la asimetría y la elasticidad en la estructura. Estos parámetros son determinados experimentalmente a partir de pruebas de laboratorio o mediante técnicas de identificación de parámetros.

Por lo cual, los parámetros del modelo de Bouc-Wen que corresponden al desplazamiento histérico (hysteretic displacement) del edificio ante fuerzas cortantes no lineal, son A , β y γ las cuales permiten controlar el tamaño y la forma en que los bucles serán generados. Mientras que $r > 1$ permite modificar la suavidad en las transiciones del lazo y la degradación de la resistencia (v) y rigidez (η) estarán en función de la duración y amplitud generada por la respuesta sísmica, obteniendo así la fuerza de histéresis (z)

que presenta la estructura.

$$\rho_i = ak_i x_i - (1 - \alpha) Dk_i z_i \tag{3.40}$$

$$= \frac{A_i \dot{x}_i - v_i (\beta_i |\dot{x}_i| z_i^{r-1} z_i - \gamma_i \dot{x}_i |z_i|^r)}{\eta_i} \tag{3.41}$$

En este sentido, el modelo de Bouc-Wen se caracteriza por su capacidad para capturar los efectos del amortiguamiento y la rigidez no lineal en la respuesta estructural. A diferencia de otros modelos que asumen una relación lineal entre la fuerza y el desplazamiento, el modelo de histéresis de Bouc-Wen tiene en cuenta el fenómeno de histéresis, que es una característica fundamental de los materiales no lineales.

Por lo que se puede acotar los modelos de Bouc-Wen con estabilidad de tipo BIBO como se presentan en la tabla 3.1, donde presentan los posibles valores que puede tomar el conjunto de variables del modelo. Donde la Clase I no solo es estable, sino también compatible con el tipo BIBO, permitiendo ser usada en sistemas reales.

Caso	Ω	Cota superior	Clase
$A > 0, \beta + \gamma > 0$ y $\beta - \gamma \geq 0$	R	$\max(z(0) , z_0)$	I
$\beta - \gamma > 0$ y $\beta - \gamma \geq 0$	$[-z_1, z_1]$	$\max(z(0) , z_1)$	II
$A < 0, \beta - \gamma > 0$ y $\beta + \gamma \geq 0$	R	$\max(z(0) , z_1)$	III
$\beta - \gamma > 0$ y $\beta + \gamma \geq 0$	$[-z_1, z_1]$	$\max(z(0) , z_1)$	IV
$A = 0, \beta + \gamma > 0$ y $\beta - \gamma \geq 0$	R	$ z(0) $	V
El resto de casos	0		0

Tabla 3.1: Clasificación de modelos Bouc-Wen BIBO estables

Por lo que la principal ventaja del modelo de histéresis de Bouc-Wen es su capacidad para describir de manera precisa y flexible una amplia gama de comportamientos no lineales. Estos comportamientos incluyen la rigidez no lineal, el amortiguamiento variable y la respuesta asimétrica ante cargas positivas.

De manera que este modelo permite simular de manera más precisa el comportamiento de las estructuras en condiciones reales, debido a esto se utiliza en problemas como: problemas de ingeniería, incluidos sistemas de múltiples grados de libertad (MDOF), edificios, marcos, respuesta bidireccional y torsional de sistemas histeréticos, continuos bidimensionales y tridimensionales, licuefacción del suelo y aislamiento básico sistemas; sus variantes y extensiones se han utilizado en control estructural.

Por lo cual, es importante entender la información que proporciona los ciclos de histéresis, en el caso de los edificios con daño estructural, es un fenómeno complejo que requiere comprensión y análisis detallados, ya que permite obtener información de las deformaciones y fisuras que se producen en los elementos estructurales durante actividad sísmica. Las cuales modifican la respuesta estructural y generan cambios en el comportamiento de la estructura.

En este sentido, los ciclos de histéresis en edificios con daño estructural proporcionan información valiosa sobre el desempeño de la estructura y su capacidad para resistir esfuerzos sísmicos. Debido a que el comportamiento de los edificios durante un evento sísmico es crucial para poder garantizar la seguridad de las personas y la integridad de las estructuras.

En este sentido, para comprender adecuadamente los ciclos de histéresis en edificios con daño estructural, es necesario tener una comprensión básica de la histéresis en sí misma. La histéresis es un fenómeno en el que la respuesta de una estructura no es completamente reversible cuando se le aplica una carga cíclica. Es decir, cuando una estructura experimenta fuerzas repetitivas, no vuelve exactamente a su estado original después de que la fuerza sea removida.

De modo que, la curva de histéresis presenta el grado de deterioro que se presenta en relación de la deformación-carga aplicada en los materiales y para los elementos estructurales es la degradación de rigidez, permitiendo una deformación plástica de manera sostenible sin romperse de la disipación de energía sísmica.

Capítulo 4

Modelo implementado con Red Neuronal Multicapa

Actualmente, para realizar el diagnóstico de daño estructural en edificios, es de vital importancia el uso de técnicas avanzadas, por lo cual las redes neuronales se han convertido en una herramienta imprescindible para ingenieros y profesionistas en la rama. Una de estas técnicas es la red neuronal multicapa o perceptrón multicapa, conocido en inglés como Multilayer Perceptron (MLP), la cual permite realizar un diagnóstico preciso y eficiente, que garantice la seguridad y la integridad de las estructuras civiles.

La red neuronal multicapa es una red neuronal artificial que está compuesta por múltiples capas de unidades de procesamiento interconectadas. Cada una de estas unidades toma una serie de entradas y produce una salida que es transmitida a la siguiente capa. Esta estructura en capas permite al perceptron aprender y reconocer patrones complejos a partir de los datos de entrada.

Este modelo de red neuronal es especialmente útil para el diagnóstico de daño estructural, debido a que permite analizar la información recopilada durante el monitoreo de las estructuras y detectar posibles anomalías o daños. Esto se debe a su capacidad para aprender y reconocer patrones que indiquen la presencia de daño, como cambios en: la rigidez, la deformación o vibraciones anormales.

En conclusión, la aplicación de la red neuronal multicapa en el diagnóstico de daño estructural se basa en la recopilación y procesamiento de datos de diferentes instrumentos de medición que monitoreen la estructura en cuestión. Estos sensores pueden medir diferentes variables, como: deformación, aceleración o la vibración de la estructura, y las variables a medir dependerán de las mediciones disponibles, de acuerdo a la accesibilidad o capacidad de obtener los instrumentos necesarios. Los datos recolectados forman la base de datos que se utilizarán en la red neuronal, ya que los procesará y analizará para identificar patrones relacionados con el daño.

4.1. Estructura de una neurona artificial

Una red neuronal es un conjunto de neuronas interconectadas entre sí, que trabajan de manera conjunta. Por lo que es necesario tener claros los conceptos asociados a cada neurona, la cual se encuentra compuesta de tres partes interconectadas: función de entrada, función de activación y función de salida, ver figura 4.1 [33].

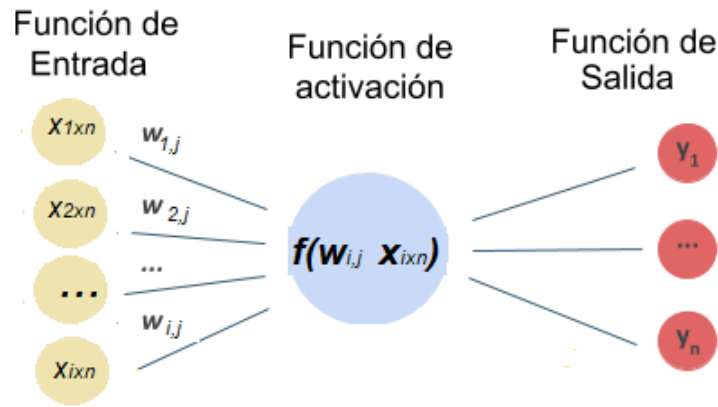


Figura 4.1: Estructura de una neurona artificial.

4.1.1. Función de entrada

La función de entrada se encuentra compuesta por tres elementos, los cuales son:

1. Valores de entrada.
2. Pesos sinópticos.
3. Regla de propagación.

Valores de entrada

Los valores de entradas y salidas en una neurona pueden ser clasificados en dos grupos:

- Binari (digital), admiten dos valores posibles que consisten en los siguientes dos alfabetos 0,1 o -1,1.
- Continua (analógica), las cuales admiten valores dentro de un determinado rango, las cuales dependen del modelo de aplicación.

Enfocándonos en este caso en específico los valores de entrada, representarán la variación de las frecuencias de anti-resonancia para obtener el comportamiento del modelo, cuando la salud estructural no se encuentra comprometida (sin daño) o cuando presente algún daño, ante una excitación sísmica. Por lo cual la entrada para una neurona de diagnóstico se denotan por:

$$x_{i \times n}(t) \quad (4.1)$$

donde i representara el valor de frecuencia en un tiempo específico y n el número del piso en el que se obtuvo la medición.

Peso sinóptico

El peso sinóptico [71] representa la intensidad de la conexión sinóptica entre dos neuronas, entre cada neurona presináptica i y la neurona postsináptica j . Los pesos sinópticos se denotan por:

$$w_{i,j} \quad (4.2)$$

De modo que los pesos sinópticos serán determinados por medio del entrenamiento de la red neuronal propuesta, a partir del uso de los datos obtenidos por el prototipo tipo edificio.

Definiendo así la intensidad de interacción entre la neurona presináptica i y la postsináptica j . Lo que nos lleva a los dos tipos de sinapsis:

- La sinapsis excitadora se presenta cuando el peso sinóptico tiene un valor positivo, por lo cual excitara a la neurona postsináptica.
- La sinapsis inhibidora se presenta cuando el peso sinóptico tiene un valor negativo, por lo cual inhibirla a la neurona postsináptica.

Regla de propagación

La regla de propagación ($h_i(t)$) [72] determina el valor de la neurona postsináptico, resultante de la interacción de la neurona i en función de sus pesos y entradas. Es decir:

$$h_i(t) = (w_{i,j}, x_{i \times n}(t)) \quad (4.3)$$

Las reglas de propagación más utilizadas en los diferentes modelos de redes neuronales son:

- Sumatoria de las entradas pesadas, es la suma de los productos escalares de todos los valores de entrada por sus correspondientes pesos sinópticos.

$$net_i(t) = h_i(t) = \sum_{j=1}^n (w_{i,j} x_{i,j}(t)), \text{ con } j= 1, 2, \dots, n \quad (4.4)$$

- Productoria de las entradas pesadas, es el producto ponderado de todos los valores de entrada multiplicados por sus correspondientes pesos sinópticos.

$$net_i(t) = h_i(t) = \prod_{j=1}^n (w_{i,j} x_{i,j}(t)), \text{ con } j= 1, 2, \dots, n \quad (4.5)$$

- Máximo de las entradas pesadas, solo se toma en consideración el valor de entrada más fuerte y este es previamente multiplicado por su peso sinóptico correspondiente.

$$net_i(t) = h_i(t) = \max_{j=1}^n (w_{i,j} \times x_{i,j}(t) x_{i,j}(t)), \text{ con } j= 1, 2, \dots, n \quad (4.6)$$

4.1.2. Función de activación

La versatilidad de las redes neuronales se presenta por el uso de funciones de activación [73], las cuales determinan la salida que se obtiene de una neurona dada una entrada o conjunto de entradas. Por lo cual, la función de activación representa la relación entre el estado de activación actual ($a_i(t)$), con base a su estado anterior ($a_i(t - 1)$) y de su potencial postsináptico actual. Como se presenta en la siguiente ecuación:

$$a_i(t) = f_i(a_i(t - 1), h_i(t)) \quad (4.7)$$

o

$$a_i(t) = f_i(a_i(t - 1), net_i(t))$$

donde, Sin embargo, en algunos modelos de redes neuronales, se suele considerarse que el estado actual de la neurona no depende de su estado anterior, sino únicamente del actual, por lo que se reescribe como:

$$\begin{aligned} a_i(t) &= f_i(h_i(t)) \\ & \text{o} \\ a_i(t) &= f_i(\text{net}_i(t)) \end{aligned} \tag{4.8}$$

Además, es importante considerar que la función de activación debe satisfacer una característica esencial: tener un comportamiento no lineal, ya que esto permite una mejor adaptación del modelo, dado que se puede utilizar una base de datos con un mayor número de mediciones.

Por lo que se busca que la función de activación al momento de obtener su derivada sea simple, ya que esto permite minimizar el coste computacional y poder así tener un conjunto de valores de salida dentro de un rango determinado como $(0, 1)$ o $(-1, 1)$.

Función sigmoide

La función sigmoide o logística se puede observar gráficamente en la figura 4.2, esta función asigna los valores de entrada en un rango de salida que se encuentra entre cero y uno, donde los valores altos tienden de manera asintótica a 1 y mientras que los valores muy bajos tienden de manera asintótica a 0. A continuación se presentan las características de esta función de activación:

- La función sigmoide no se encuentra centrada en cero, es decir, no es simétrica en cero. En consecuencia, se dificulta el aprendizaje y validación, debido a que el gradiente se desplaza hacia una dirección particular.
- Presenta lenta convergencia, es decir, requiere un periodo de tiempo largo para que la red neuronal converja a los valores deseados.
- Este tipo de función es empleada principalmente empleada en la última capa de una red neuronal debido a que se utiliza para la clasificación binaria.
- Presenta el problema de saturación, el cual depende de la función de activación debido a que sí es muy plana en los extremos. Al momento de derivar la función en esos puntos se obtendrán valores muy pequeño, por lo tanto, el gradiente será cercano a cero.
- Presenta el problema de gradiente desvaneciente, es decir, que a medida que se realiza la retropropagación el gradiente se va haciendo más pequeño al punto que la actualización de los pesos sinópticos tienda a cero.
- Presenta el problema de explosión de gradientes, es decir, el gradiente es demasiado grande, haciendo que la red neuronal se vuelva inestable. Por lo cual, los pesos sinópticos crecerán desmesuradamente al punto de tener valores *NaN*.

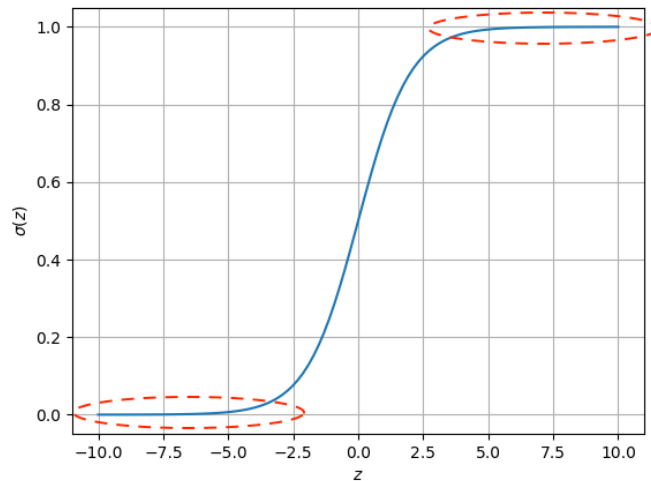


Figura 4.2: Función sigmoide.

Función tangente hiperbólica

La función tangente hiperbólica es una función no lineal, como se puede observar gráficamente en la figura 4.3. Debido a esto, las redes neuronales pueden modelar relaciones no lineales entre los datos.

Donde los valores de salida se encuentran dentro del rango de $(-1, 1)$, por lo tanto, los valores negativos tienden de manera asintótica a -1 y los positivos tienden de manera asintótica a 1 . A continuación se presentan las características de esta función de activación:

- Presenta el problema de saturación y de gradiente desvaneciente.
- Es comúnmente empleada en redes neuronales recurrentes o para clasificación binaria.
- La salida de la función se encuentra centrada en cero.
- Presenta un error en la etapa de aprendizaje al momento de utilizar el algoritmo de propagación hacia atrás.
- Presenta lenta convergencia.

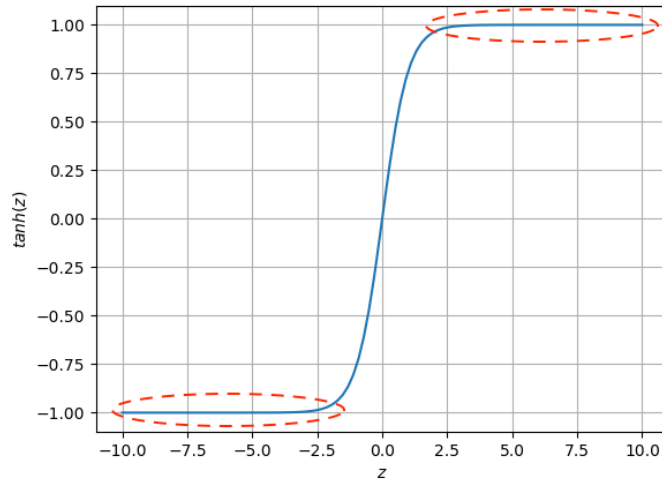


Figura 4.3: Función tangente hiperbólica.

Función unidad lineal rectificada - "ReLU"

La función unidad lineal rectificada, conocida en inglés como; Rectified Lineal Unit (ReLU) se puede observar gráficamente en la figura 4.4, es considerada la función de activación que más se asemeja al funcionamiento de una neurona biológica. De modo que los valores de entrada menores o iguales a cero se obtiene una salida igual a cero, de lo contrario se obtiene que la salida es igual a su entrada. A continuación se presentan las características de esta función de activación:

- Esta función de activación se emplea para redes neuronales de arquitectura profunda o redes neuronales convolucionales, debido a que permite un rápido aprendizaje.
- No presenta el problema del gradiente desvaneciente.
- Presenta Dying ReLU, es decir, tiende a volver inactivas el 60% de las neuronas debido a que presenta una gran capacidad de aprendizaje, reduciendo así la capacidad de aprendizaje del modelo si se deja entrenando por un largo periodo de tiempo.
- Presenta la activación de Sparse, es decir, solo se activa la función si los valores son positivo.
- Se presenta la muerte de neuronas, es decir, la función de activación da como resultado cero y al momento de derivarla también es cero.

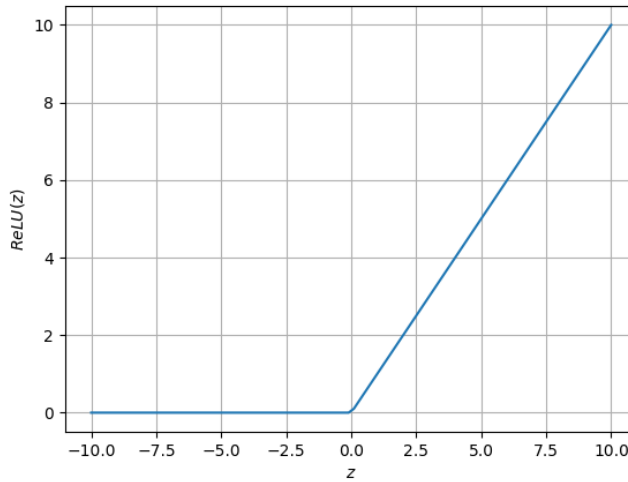


Figura 4.4: Función ReLU.

Función unidad lineal rectificada - "Leaky - ReLU"

La función "Leaky - ReLU" se puede observar gráficamente en la figura 4.5, ante la entrada de valores negativos estos son multiplicados por un coeficiente rectificativo, mientras que al presentarse valores mayores o iguales a cero estos se mantienen según entran. A continuación se presentan las características de esta función de activación:

- Evita la muerte de neuronas debido a la pendiente que existe cuando $x < 0$.
- Esta función de activación es principalmente usada en redes neuronales convolucionales o para redes de identificación en imágenes.

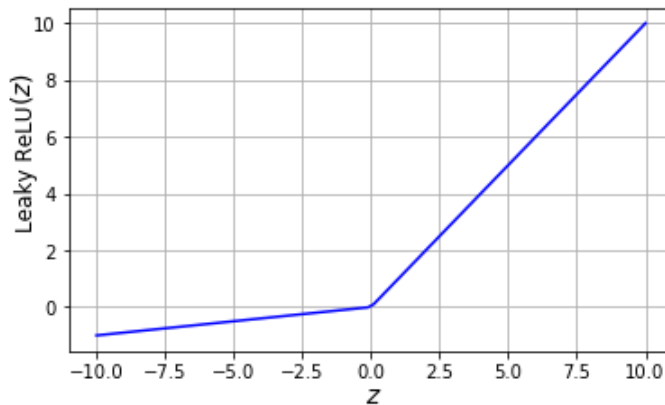


Figura 4.5: Función Leaky ReLU.

Función unidad lineal rectificada - "Softmax"

La función "Softmax" se puede observar gráficamente en la figura 4.6, esta función determina la probabilidad que tiene un valor de entrada de pertenecer a una clase determinada, A continuación se presentan las características de esta función de activación:

- Utilizada en modelos de regresión logística para clasificación múltiple.
- Presenta un buen rendimiento cuando se emplea en las últimas capas.
- Empleada cuando se requiere obtener una representación en forma de probabilidades.
- Se emplea para la clasificación de datos, siendo un ejemplo de esto que ante la entrada de una imagen de frutas se pretende determinar el tipo de fruta al que pertenece, entregando a la salida la probabilidad, siendo el resultado la clase que tenga mayor probabilidad, donde la suma de las probabilidades debe ser igual a 1.

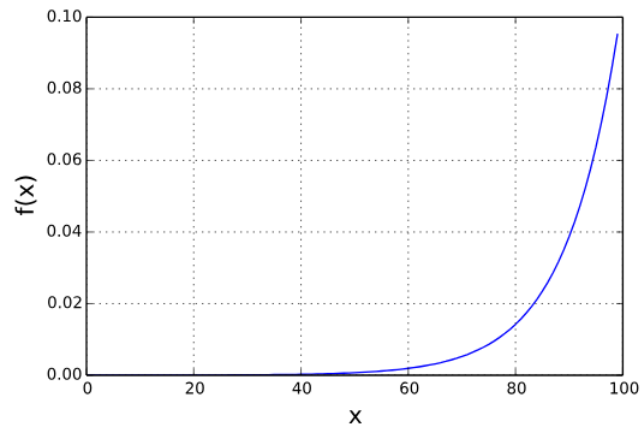


Figura 4.6: Función Softmax.

En la tabla 4.1, se presentan las funciones de activación anteriormente mencionadas denotando su nombre, función y rango de respuesta.

	Función	Rango
Sigmoide	$f(x) = \frac{1}{1+e^{-x}}$	$[0,+1]$
Tangente hiperbólica	$f(x) = \frac{2}{1+e^{-2x}} - 1$	$[-1,+1]$
ReLU	$f(x) = \begin{cases} 0 & \text{para } x < 0 \\ x & \text{para } x \geq 0 \end{cases}$	No está acotada
Leaky ReLU	$f(x) = \begin{cases} 0 & \text{para } x < 0 \\ a \cdot x & \text{para } x \geq 0 \end{cases}$	No está acotada
Softmax	$f(z)_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}$	$[0,+1]$

Tabla 4.1: Funciones de activación.

4.1.3. Función de salida

La función de salida [74], se encuentra compuesta por una función de activación, donde si el valor de entrada está por debajo de un umbral determinado, no se obtiene un valor de salida para la neurona subsiguiente, puesto que, no cualquier valor es permitido como entrada, debido a que los valores de salida están comprendidos por un rango determinado con base a la función de activación.

En resumen, el valor de la función de salida será transferido a la siguiente capa de neuronas

vinculas o si nos encontramos en la última capa de neuronas serán los valores que determinen la identificación o clasificación de los datos de entrada. Las funciones de salida más comunes:

- Función identidad, se presta cuando los valores que entran a la neurona son idénticos a los valores de salida.
- Binaria, $\begin{cases} 1 & \text{si } f_i(t) \geq \xi_i \\ 0 & \text{de lo contrario} \end{cases}$, donde ξ_i es el umbral.

La función de salida es dentada por:

$$y_i(t) = f_i(h_i(t)) \tag{4.9}$$

o

$$y_i(t) = f_i(net_i(t))$$

En este caso, si nos enfocamos en diseñar una red neuronal que identifique el daño que presenta una estructura en cada uno, de sus pisos, la identificación de daño estructural será a partir de la pérdida de rigidez que presente el prototipo. Esto será por medio de la comparación de los factores que representen la razón entre la variación de la rigidez y la rigidez inicial.

Por lo cual es conveniente elegir una función de activación que permita clasificar los datos de la siguiente manera:

- Para indicar que el piso se encuentra sin daño, la salida en la red neuronal será igual a $y_i = 0$.
- Para identificar la presencia de daño, la salida en la red neurona debe ser $0 < y_i \leq 1$.

4.2. Estructura de una red neuronal multicapa

Las redes neuronales multicapa son modelos, que se encuentran contruidos a partir de un número determinado de neuronas formando niveles o capas, como se muestra en la figura 4.7 [33]. Donde las partes principales de una red neuronal multicapa son [75]:

1. Capa de entrada.
2. Capas ocultas.
3. Capa de salida.

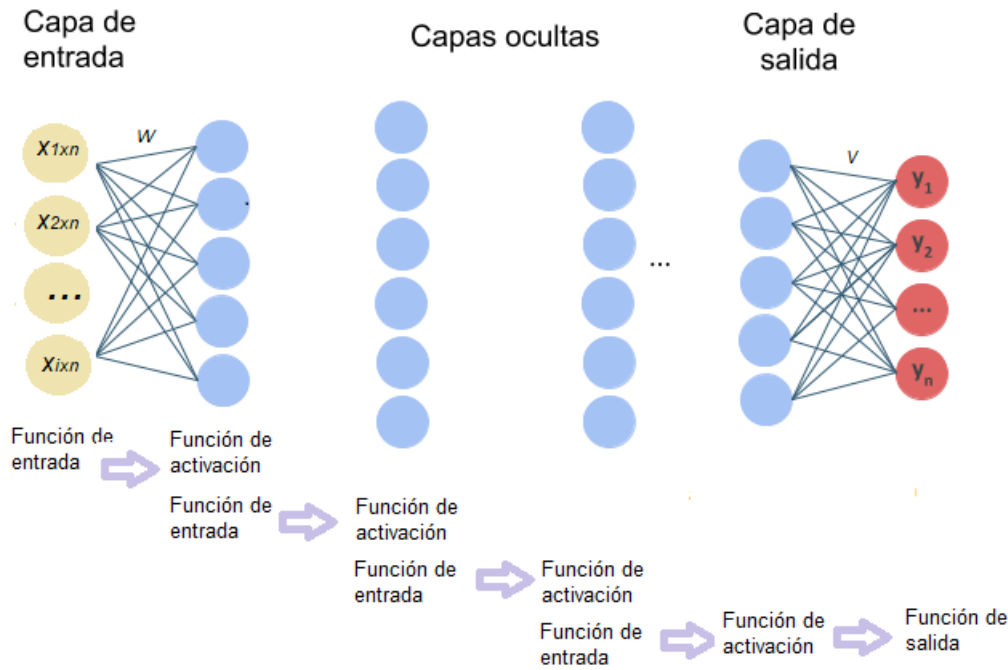


Figura 4.7: Red neuronal multicapa totalmente interconectada.

4.2.1. Capa de entrada

Es una capa transparente, es decir, en una capa de neuronas se recibe los valores de entrada, donde los datos son: analizados o clasificados, para así pasar la información que se quiere a la siguiente capa. En este caso, la capa de entrada tiene como base de datos las frecuencias a las que vibra el prototipo ante una excitación sísmica.

4.2.2. Capas ocultas

Son capas que se encuentran interconectadas dentro de la red neuronal, las cuales no tienen contacto directo con la salida, estas pueden estar o no presentes, debido a que dependen de: la selección de hiperparámetros elegidos y de la complejidad de los patrones que se desean aprender.

Cuando la red está conformada por una única capa de neuronas, se conoce como red neuronal monocapa, donde las neuronas que conforman la capa entrada cumplen la función de entrada y salida simultáneamente.

Por otra parte, cuando la red está compuesta por más de una capas, se habla de redes neuronales multicapa, donde cada capa oculta analiza la salida de la capa anterior, con la finalidad de proporcionar un mejor aprendizaje, hasta llegar a la última capa oculta.

4.2.3. Capa de salida

Esta capa proporciona los índices de daño que se presentan en cada piso del prototipo hacia el exterior, siendo este el procesamiento de datos que realiza la red neuronal, de acuerdo a su aprendizaje.

4.2.4. Formas de conexión entre neuronas

La conectividad entre la red neuronal está relacionada con la forma en que la señal de salida de una neurona pueden ser una entrada de otras neuronas o incluso ser una entrada de sí mismas (conexión auto-recurrente), teniendo dos tipos de conexiones:

- Propagación hacia delante o redes feedforward: cuando las interconexiones entre las distintas neuronas, ninguna salida de las neuronas es entrada de neuronas del mismo nivel o de niveles precedentes, por lo que siguen un único sentido, es decir, se realiza la propagación desde la entrada de la red hacia su salida.
- Propagación hacia atrás o redes feedbackward: cuando las interconexiones entre distintas salidas pueden ser conectadas como entradas de neuronas de su mismo nivel o de niveles previos o, incluso, de sí mismas, como es el caso de las redes neuronales recurrente.

4.2.5. Hiperparámetros

Los hiperparámetros o arquitectura de las redes consiste en la organización y disposición de las neuronas en la red, formando capas. Donde los parámetros fundamentales de la red son:

- Número de capas.
- Número de neuronas por capa.
- Grado de conectividad.
 - Totalmente conectadas, es decir, todas las neuronas de una capa de entrada se encuentran conectadas con todas las neuronas de la siguiente capa.
 - Parcialmente conectadas, es decir, no todas las neuronas de entrada están conectadas a las neuronas de la siguiente capa.
- Forma de conexión entre neuronas.
- Número de interacciones o margen de error aceptable.
- Valor de los pesos sinópticos iniciales.

4.3. Red Neuronal Multicapa

Este tipo de red neuronal se encuentra compuestas por múltiples capas de neuronas interconectadas [76], donde cada capa contiene un conjunto de neuronas que trabajan en conjunto para procesar la información. Donde la primera capa es la capa de entrada, la cual recibe los las frecuencias anti-resonancia, que en este caso son obtenidas a partir del monitorio del prototipo tipo edificio ante la señal de excitación sísmica a los cual estará sujeto el modelo estructural y los envía a la siguiente capa para su procesamiento. Las capas ocultas, situadas entre la capa de entrada y la capa de salida, son responsables de realizar cálculos más complejos y extraer características clave de los datos. Finalmente, la capa de salida entregará los indicadores de daño encontrado en cada piso de la estructura.

Así mismo, cada capa puede tener un número diferente de neuronas e incluso una función

de activación diferente, por lo que se emplea la utilización de superíndices para identificar el número de capas; por lo tanto, se agrupan los pesos sinópticos de la primera capa en una matriz que es denotada por W^1 y la segunda matriz que englobe los pesos sinópticos es W^2 y así sucesivamente.

4.3.1. Aplicaciones de redes neuronales multicapa

Las redes neuronales multicapa tiene diferentes aplicaciones, pero las principales están enfocadas a:

Aproximación de funciones

Al utilizar una red neuronal para aproximación de funciones se puede generalizar diferentes modelos en uno solo y así predecir la respuesta que se deberá obtener partir de una base de datos desconocidos, basándose de los aprendizajes pasados. Principalmente, se emplean en los sistemas de control, debido a que permite encontrar una función de retroalimentación adecuada que mapee las salidas medidas hasta las entradas de control.

Ejemplo 5. Identificación de sistema lineal con redes neuronales

$$y(k) = -0,12y(k-1) + 0,7y(k-2) + u(k) \quad (4.10)$$

Primero se utilizó una red neuronal monocapa (ver figura 4.8) para realizar una identificación del sistema presentado en el ejemplo 5, en este caso se tiene un aprendizaje y prueba que se podría considerar bueno, pero no identifica el sistema al 100%. Por lo cual presenta un error considerable y si se tratara de un sistema que requiere una exactitud precisa, este modelo de red neuronal no sería una opción recomendable para su implementación.

Por otra parte, si se emplea una red neuronal multicapa (ver figura 4.9), se presenta un error de identificación casi imperceptible a comparación de una red neuronal monocapa, haciendo que si se requiera más precisión se modificarían los hiperparámetros.

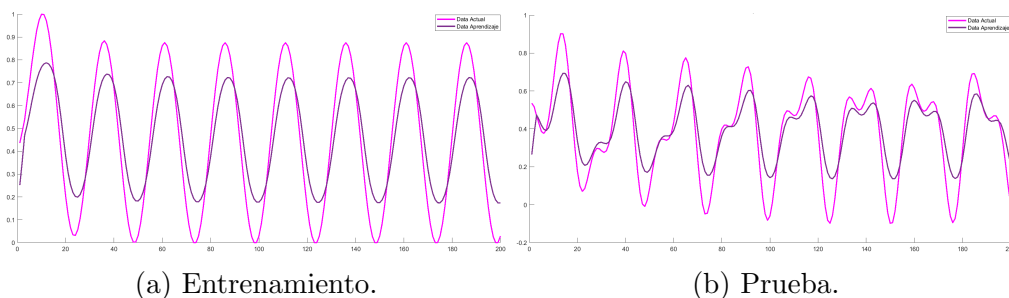


Figura 4.8: Identificación de sistema 4.10, con red neuronal monocapa.

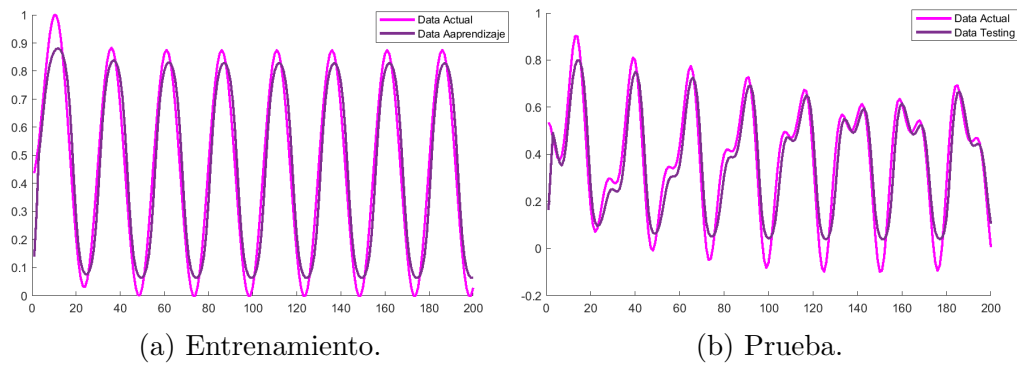


Figura 4.9: Identificación de sistema 4.10, con red neuronal multicapa.

De manera que al utilizar las redes neuronales con un enfoque de identificación de sistemas, se utilizara de una base de datos que contenga las frecuencias a las que vibre el prototipo ante una excitación sísmica. Para poder así obtener el comportamiento de la energía cuando se presente daño estructural en el prototipo.

Clasificación de patrones

Por otra parte, si se utilizan las redes neuronales multicapa, con un enfoque de clasificación de patrones, ya sea de una función binaria como la función "XOR" o bien de una base de datos con decenas de características. De modo que para realizar una clasificación de datos, se requiere una fase de aprendizaje, en la cual recibe un conjunto de datos de entrada junto con las etiquetas de clase correspondientes.

Y a medida que la red procesa los datos, se envía una señal de retroalimentación desde la capa de salida hacia las capas ocultas, permitiendo así un ajustar en los parámetros de manera automáticamente. Por lo tanto, si se requiere minimizar la diferencia entre las salidas esperadas y las salidas obtenidas, y así mejorar la probabilidad de pertenencia a la clase correcta.

4.3.2. Aprendizaje por algoritmo de retropropagación (Backpropagation)

El aprendizaje por algoritmo de retropropagación o propagación hacia atrás de errores, es un método que permite calcular el gradiente a partir de un proceso de propagación-adaptación. Es decir, los datos de entrada se propaga capa por capa hasta llegar a la salida de la red, en donde, la señal de salida es comparada con la salida deseada, obteniendo así una señal de error para cada una de las salidas.

A partir de tener las salidas de error se realiza una propagan hacia atrás capa por capa, es decir, desde la capa de salida hasta a la primera capa de neuronas. Donde se repetirá un proceso de ponderación aproximado de la contribución relativa que tuvo cada neurona de la red con respecto a la señal de salida, hasta que todas las neuronas de la red hayan recibido una señal de error que describa su contribución al error total.

La importancia de este tipo de aprendizaje radica en que a medida que se entrena la red neuronal, las capas ocultas aprenden a reconocer distintas características de los valores de entrada. De modo que el entrenamiento permite que las capas ocultas responderán a patrones arbitrarios de entrada que contenga ruido, siempre y cuando se asemeje a aquella característica que las neuronas individuales hayan identificado.

Las ecuaciones que describen este algoritmo son:

$$a^{m+1} = f^{m+1} (W^{m+1} a^m + b^{m+1}) \quad \text{Para } m=0, 1, \dots, m-1 \quad (4.11)$$

Donde, m es el número de capas que tiene la red neuronal, b^{m+1} son los pesos sinópticos de las conexiones entre las neuronas, a^{m+1} las de la neurona dentro de las capas ocultas y $x_{i \times n}$ son los valores de entrada externos. Por lo que, las neuronas de la primera capa reciben entradas externas:

$$a^0 = x_{i \times n} \quad (4.12)$$

La última capa de neuronas proporciona la salida de la red neuronal, la cual se encuentra descrita por:

$$a = a^m \quad (4.13)$$

la cual proporciona el punto de partida para la ecuación 4.11. Por ejemplo, para una red de 3 capas es descrita como:

$$a^3 = f^3 (W^3 f^2 (W^2 f^1 (W^1 x_{i \times n} + b^1) + b^2) + b^3) \quad (4.14)$$

Índice de desempeño

Los índices de desempeño o indicadores de rendimiento, están directamente relacionados con medir la diferencia entre los valores previstos por el modelo predictivo y los valores reales. Proporcionando así una estimación de la capacidad que posee el modelo predictivo puede predecir el valor deseado.

En este sentido, los índices de desempeño en las redes neuronales evalúan el aprendizaje que esta tiene, por esta razón se define el error medio cuadrático (MSE), el cual describe la relación de este método con la regresión lineal, obteniendo así un valor numérico que exprese el desempeño de la red. Lo que nos lleva a que el error medio cuadrático es definido por:

$$MSE = \frac{1}{2} \sum_{k=1}^n [y(k) - \hat{y}(k)]^2 \quad (4.15)$$

donde, $y(k)$ y $\hat{y}(k)$ k , es la salida deseada y estima en una interacción k respectivamente. Por lo que si trasladamos este método para desarrollar un algoritmo que ajuste los parámetros de la red para minimizar error, se requiere un conjunto de ejemplos del comportamiento adecuado que debe tener la red neuronal:

$$(x_1, y_1), (x_2, y_2), \dots (x_Q, y_Q) \quad (4.16)$$

donde, x son los valores de entrada a la red y y , es la salida correspondiente, ya que a medida que cada entrada se aplica a la red, la salida, se compara con el valor deseado, lo que nos lleva a:

$$F(\theta) = E [e^2] = E [(y - a)^2] \quad (4.17)$$

donde, θ será el vector que almacenara el cuadrado del sesgo de las predicciones de la red y si la red tiene múltiples salidas, esto se describe como:

$$F(\theta) = E [e^T e] = E [(y - a)^T (y - a)] \quad (4.18)$$

Pero si aproximamos el error medio cuadrático que se tiene por cada interacción k , se obtiene:

$$\hat{F}(x) = (t(k) - a(k))^T (t(k) - a(k)) = e^T(k) e(k) \quad (4.19)$$

De modo que si utilizamos este algoritmo para un aprendizaje profundo, se usa el gradiente descendiente con la finalidad de actualizar los pesos de cada una de las neuronas, desacuando a su contribución aproximada al error total, se tiene:

$$w_{i,j}^m(k+1) = w_{i,j}^m(k) - \alpha \frac{\partial \hat{F}}{\partial w_{i,j}^m} \quad (4.20)$$

$$b_i^m(k+1) = b_i^m - \alpha \frac{\partial \hat{F}}{\partial b_i^m} \quad (4.21)$$

donde, α es un valor constante en algoritmo de gradiente descendiente que representa la tasa de aprendizaje. La cual define la velocidad a la que el algoritmo alcanza los pesos sinópticos óptimos, de modo que , si la tasa de aprendizaje muy alta, esto significaría que no se podrían alcázar, por otra parte, si la tasa de aprendizaje es muy pequeña esto significará que se requiere de un mayor número de interacciones para alcanzar los pesos sinópticos óptimos.

Por lo que, si utilizamos este índice de rendimiento, se podrá obtener los pesos sinópticos óptimos para cada neurona, teniendo así un aprendizaje con el menor error posible. Y poder así estimar el parámetro z que representa la fuerza de histéresis en el modelo Bouc-Wen, de modo que se podrá determinar la manera en que se disipa la energía, una estructura ante una excitación sísmica.

Regla de la cadena

La principal característica de las redes neuronales es poder aprender y adaptarse a partir de un conjunto de frecuencias anti-resonancia, por lo cual son capaces de realizar predicciones sobre el comportamiento de la disipación de energía que presente el prototipo ante excitación sísmica. Esto se logra a través del aprendizaje supervisado, donde la red aprende con datos obtenidos cuando el prototipo se encuentra en condiciones sin daño y con daño.

En consecuencia, se requiere utilizar la regla de la cadena para poder calcular el error por medio del gradiente descendiente, esto permitirá realizar ajustes más preciso en los pesos sinópticos de cada neurona que conforme la red neuronal, teniendo así un mejor diagnóstico de daño para cada piso del prototipo.

La regla de la cadena surge de tener una función explícita f en función de una variable n , pero se requiere derivar la función f respecto a una tercera variable, es decir:

$$\frac{df(n(w))}{dw} = \frac{df(n)}{dn} \times \frac{dn(w)}{dw} \quad (4.22)$$

Ya que el error es una función indirecta de los pesos sinópticos que se encuentran en las capas ocultas, se usa la regla de la cadena para calcular las derivadas de las ecuaciones 4.20 y 4.21:

$$\frac{\partial \hat{F}}{\partial w_{i,j}^m} = \frac{\partial \hat{F}}{\partial n_i^m} \times \frac{\partial n_i^m}{\partial w_{i,j}^m} \quad (4.23)$$

$$\frac{\partial \hat{F}}{\partial b_i^m} = \frac{\partial \hat{F}}{\partial n_i^m} \times \frac{\partial n_i^m}{\partial b_i^m} \quad (4.24)$$

El segundo término se puede calcular, ya que se tiene el valor total que entra a la capa m , ya que \hat{F} es una función explícita en la cual se tienen los pesos sinópticos y los sesgo en esa capa:

$$n_i^m = \sum_{j=1}^{s^{m-1}} w_{i,j}^m a_j^{m-1} + b_i^m \quad (4.25)$$

Por lo tanto:

$$\frac{\partial n_i^m}{\partial w_{i,j}^m} = a_j^{m-1} \quad (4.26)$$

$$\frac{\partial n_i^m}{\partial b_i^m} = 1 \quad (4.27)$$

De modo que se define la sensibilidad (s_i^m), que tiene la función \hat{F} , respecto a los cambios que se presenten, en cada elemento de la entrada:

$$s_i^m \equiv \frac{\partial \hat{F}}{\partial n_i^m} \quad (4.28)$$

Por lo que las ecuaciones 4.23 y 4.24, se simplifican como:

$$\frac{\partial \hat{F}}{\partial w_{i,j}^m} = s_i^m a_j^{m-1} \quad (4.29)$$

$$\frac{\partial \hat{F}}{\partial b_i^m} \approx s_i^m \quad (4.30)$$

Llegando a que para cada peso sinóptico $w_{i,j}$ y b_j^m , se describe como:

$$w_{i,j}^m(k+1) = w_{i,j}^m(k) - \alpha s_i^m a_j^{m-1} \quad (4.31)$$

$$b_j^m(k+1) = b_j^m(k) - \alpha s_i^m \quad (4.32)$$

Si se reescriben en forma matricial se tiene:

$$W^m(k+1) = W^m(k) - \alpha s^m - (a^{m-1})^T \quad (4.33)$$

$$b^m(k+1) = b^m(k) - \alpha s^m \quad (4.34)$$

$$s^m = \frac{\partial \hat{F}}{\partial n^m} = \begin{bmatrix} \frac{\partial \hat{F}}{\partial n_1^m} \\ \frac{\partial \hat{F}}{\partial n_2^m} \\ \vdots \\ \frac{\partial \hat{F}}{\partial n_{s^m}^m} \end{bmatrix} \quad (4.35)$$

donde, W^m es la matrice en la que se agrupan todos los pesos sinópticos de las neuronas que conforman una capa m y b^m es la matriz en las que se agrupan los pesos sinópticos de las conexiones entre las neuronas en una capa m .

Sensibilidad en el algoritmo de retropropagación

La sensibilidad en el algoritmo de retropropagación se refiere a la capacidad que se tiene para capturar y responder a cambios de frecuencia que presente el prototipo ante una excitación sísmica por la presencia de daño en la estructura; en otras palabras, la sensibilidad determina cuánto afectan los cambios en las frecuencias de entrada a la salida de una red, incluso pequeños cambios en las frecuencias de entrada pueden tener un gran impacto en determinar si se presenta daño o no en la estructura. Por otro lado, si el algoritmo es insensible, los cambios en los datos de frecuencia pueden ignorarse o tener un efecto mínimo en el diagnóstico de daño. La sensibilidad demasiado alta puede hacer que la red sea demasiado sensible a las variaciones de frecuencia de entrada e inestable y propensa a errores. Por lo cual, una sensibilidad demasiado baja puede hacer que la red sea lenta para adaptarse a los cambios en las frecuencias de entrada y, en última instancia, menos precisa para obtener un diagnóstico que se considere útil.

La relación de recurrencia de las sensibilidades, se define a partir de la siguiente matriz jacobiana:

$$\frac{\partial n^{m+1}}{\partial n^m} = \begin{bmatrix} \frac{\partial n_1^{m+1}}{\partial n_1^m} & \frac{\partial n_1^{m+1}}{\partial n_2^m} & \dots & \frac{\partial n_1^{m+1}}{\partial n_{s^m}^m} \\ \frac{\partial n_2^{m+1}}{\partial n_1^m} & \frac{\partial n_2^{m+1}}{\partial n_2^m} & \dots & \frac{\partial n_2^{m+1}}{\partial n_{s^m}^m} \\ \vdots & \vdots & & \vdots \\ \frac{\partial n_{s^{m+1}}^{m+1}}{\partial n_1^m} & \frac{\partial n_{s^{m+1}}^{m+1}}{\partial n_2^m} & \dots & \frac{\partial n_{s^{m+1}}^{m+1}}{\partial n_{s^m}^m} \end{bmatrix} \quad (4.36)$$

Donde se consideran los pesos sinópticos $w_{i,j}$ y b_j^m , por lo tanto, la matriz se reescribe como:

$$\begin{aligned} \frac{\partial n_1^{m+1}}{\partial n_j^m} &= \frac{\partial \left(\sum_{l=1}^{s^m} w_{i,l}^{m+1} a_l^m + b_i^{m+1} \right)}{\partial n_j^m} = w_{i,j}^{m+1} \frac{\partial a_j^m}{\partial n_j^m} \\ &= w_{i,j}^{m+1} \frac{\partial f^m(n_j^m)}{\partial n_j^m} = w_{i,j}^{m+1} f^m(n_j^m) \end{aligned} \quad (4.37)$$

Donde se tiene que:

$$f^m(n_j^m) = \frac{\partial f^m(n_j^m)}{\partial n_j^m} \quad (4.38)$$

Por lo cual la matriz jacobiana se reescribe como:

$$\begin{aligned} \frac{\partial n^{m+1}}{\partial n^m} &= W^{m+1} F^m(n^m) \\ F^m(n^m) &= \begin{bmatrix} f^m(n_1^m) & 0 & \dots & 0 \\ 0 & f^m(n_2^m) & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & f^m(n_{s^m}^m) \end{bmatrix} \end{aligned} \quad (4.39)$$

Se reescribe la relación de recurrencia de las sensibilidades usando la regla de la cadena:

$$\begin{aligned} s^m &= \frac{\partial \hat{F}}{\partial n^m} = \left(\frac{\partial n^{m+1}}{\partial n^m} \right)^T \frac{\hat{F}}{\partial n^{m+1}} = \\ &= F^m(n^m) (W^{m+1})^T s^{m+1} \end{aligned} \quad (4.40)$$

Las sensibilidades se propagan hacia atrás a través de la red desde la última capa a la primera capa:

$$s^m \rightarrow s^{m-1} \rightarrow \dots \rightarrow s^2 \rightarrow s^1 \quad (4.41)$$

Obteniendo la relación de recurrencia de las sensibilidades en la ecuación 4.40, se completa el algoritmo de repropagación, por lo tanto, la salida de la última capa de la red neuronal está dada por:

$$s_i^m = \frac{\partial \hat{F}}{\partial n_i^m} = \frac{\partial (y - a)^T (y - a)}{\partial n_i^m} = \frac{\partial \sum_{j=1}^{s^m} (t_j - a_j)^2}{\partial n_i^m} = -2(y_i - a_i) \frac{\partial a_i}{\partial n_i^m} \quad (4.42)$$

$$\frac{\partial a_i}{\partial n_i^m} = \frac{\partial a_i^m}{\partial n_i^m} = \frac{\partial f^m(n_i^m)}{\partial n_i^m} = f^m(n_i^m) \quad (4.43)$$

$$s_i^m = -2(y_i - a_i) f^m(n_i^m) \quad (4.44)$$

$$s^m = -2F^m(n^m)(y - a) \quad (4.45)$$

4.4. Evaluación del modelo implementado redes neuronales

4.4.1. Prototipo de estructura tipo edificio

Los experimentos se realizarán en un prototipo de tipo edificio que permite adquirir la información del comportamiento de una estructura ante la presencia de actividad sísmica, este equipo se encuentra montado en una mesa vibratoria que es accionada por servo-motores, como se presenta en la figura 4.10 y en el Apéndice A se presentan las especificaciones de este.

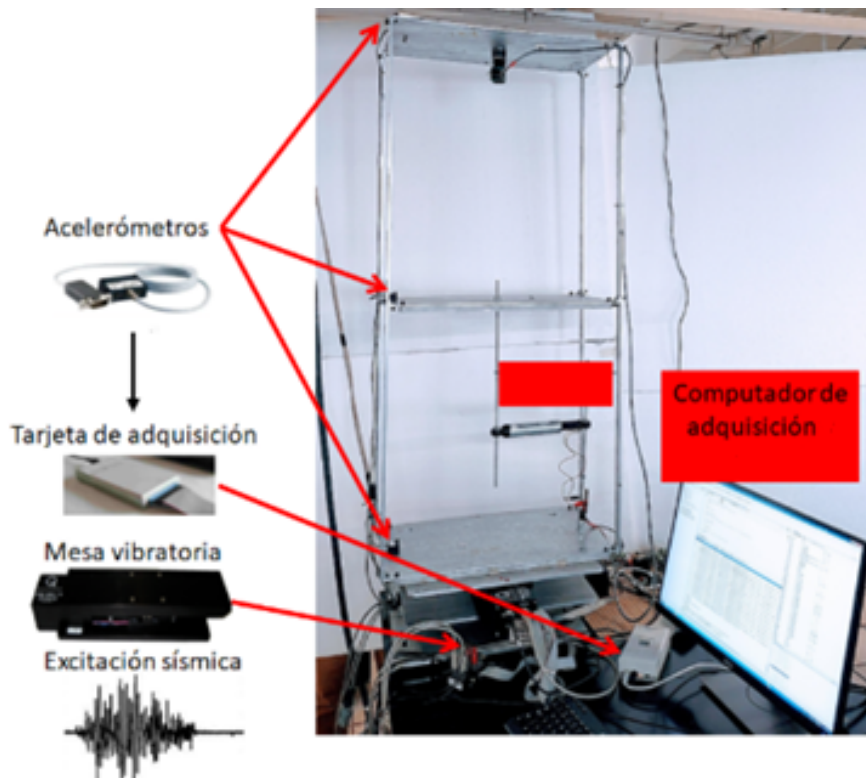


Figura 4.10: Prototipo de edificio de 2 niveles localizado en el Laboratorio DCA.

El prototipo es accionado a través de la mesa vibratoria que recibe una señal de activación por medio de una base de datos en la que se encuentran registrados los componentes del registro sísmico de Northridge. Donde ocurrió un terremoto de empuje ciego (falla que no alcanza la superficie de la Tierra), de magnitud 6,7 Richter, el día 17 de enero de 1994 a las 04:30 hora estándar del Pacífico en la región del Valle de San Fernando de la ciudad de Los Ángeles (ver figura 4.11), que ha sido escalado; por lo cual, la magnitud de desplazamiento se ajusta para su coincidencia con el prototipo del edificio.

Por lo tanto, para determinar el comportamiento del prototipo ante una excitación sísmica registrada se requiere el uso de instrumentos de medición (acelerómetros) que determinen el comportamiento de las aceleraciones, con el fin de estimar los parámetros de velocidad y el desplazamiento que presente cada uno de los pisos de la estructura

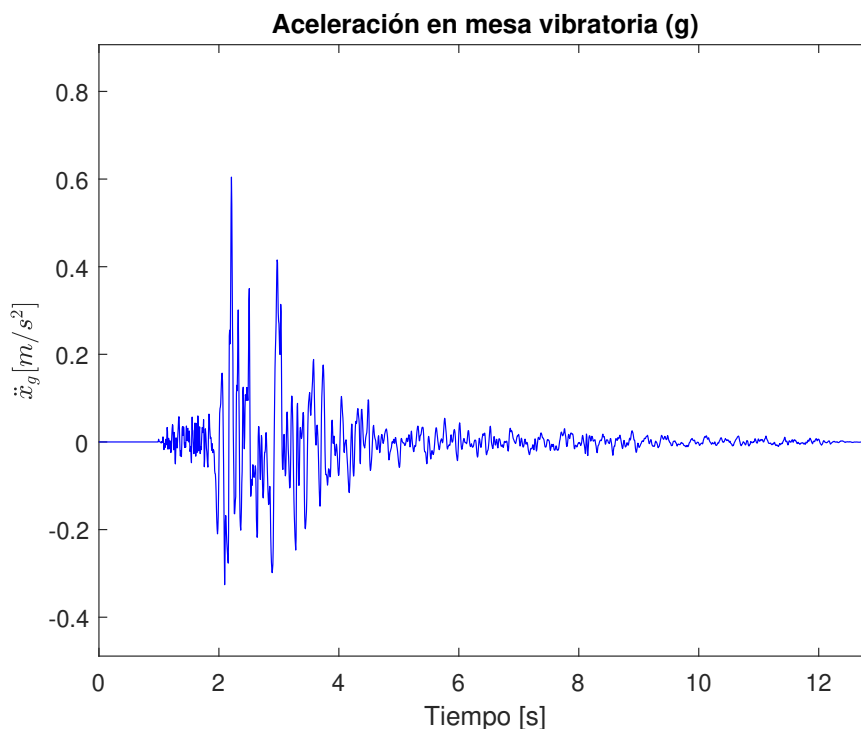


Figura 4.11: Comportamiento de las componentes del registro sísmico de Northridge 1994.

Para poder desarrollar la detección de daño es necesario tener cierta información previa sobre las condiciones nominales de la estructura de estudio, por lo que, se utiliza una adquisición de datos. Es decir, se requiere una interfaz de comunicación entre el prototipo y un ordenador.

Los instrumentos de medición, en este caso los acelerómetros, estarán conectados a una tarjeta de adquisición de datos, manteniendo así una comunicación entre el prototipo y un programa desarrollado en MATLAB 2012b/Simulink (ver figura 4.12), de modo que se adquieran las mediciones de aceleración en cada uno de los pisos de la estructura ante una excitación sísmica.

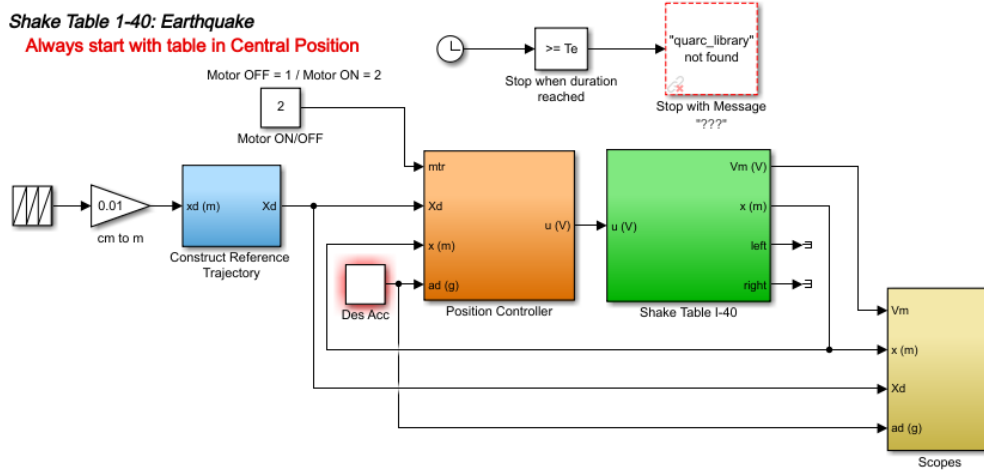


Figura 4.12: Adquisición de datos empleando MATLAB 2012b/Simulink.

4.4.2. Diseño del experimento

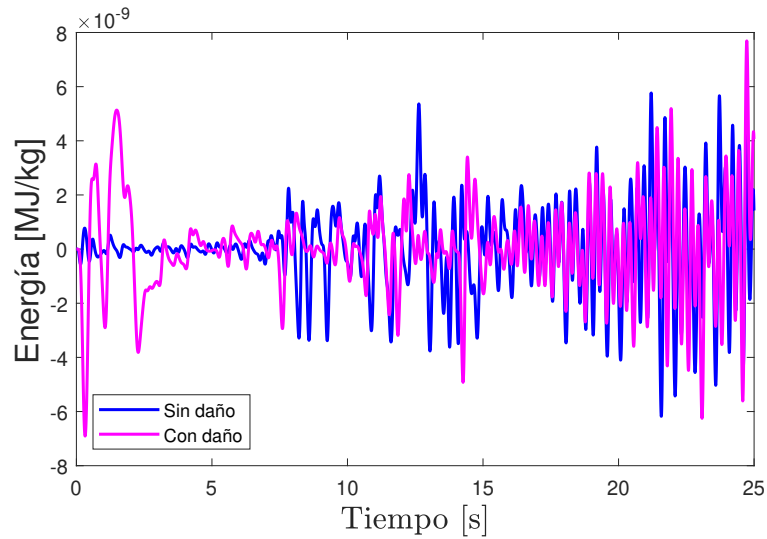
Se considera a un sismo como una fuerza de excitación externa que afecta a una estructura, al punto de causar daños. Donde la gravedad de estos, dependerán de la capacidad que tiene la estructura para disipar la energía cinética a la cual es sometida. Por lo que, en la actualidad, los reglamentos de construcción recomiendan, aprovechar la ductilidad de los materiales empleados en construcción, haciendo que la supervivencia de los edificios ante la presencia de un sismo severo dependa de esta.

Cuando la estructura entra al intervalo de comportamiento inelástico, la respuesta se reduce debido a que existe disipación de energía. Es decir, que el equilibrio energético que presenta en un sistema estructural durante un sismo se expresa como:

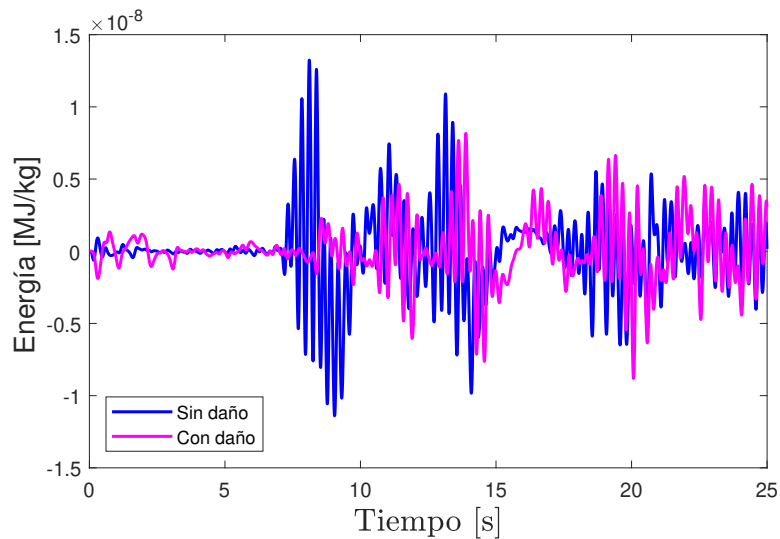
$$T + E_S + E_H + E_D = E_I \quad (4.46)$$

donde E_I es la energía total inducida por el sismo; T , E_S , E_D y E_H , son: energía cinética, energía de deformación elástica, energía disipada por el amortiguamiento viscoso de la estructura y energía por deformación de la estructura en el rango inelástico, respectivamente. La energía presente durante la vibración elástica se representa por: $T + E_S$, mientras que la energía disipada por el sistema es: $E_H + E_D$.

donde T y E_S representan un bajo porcentaje de la energía recibida durante el sismo, las cuales son: liberadas por el sismo y absorbida por la estructura.



(a) Energía en piso 1.



(b) Energía en piso 2.

Figura 4.13: Energías con presencia de daño en el piso 2.

De modo que, la energía que es disipada por medio del amortiguamiento a través del movimiento relativo de las partes que componen los edificios, pudiendo ser estas estructurales o no; y también representan un bajo porcentaje de la energía recibida. La energía disipada en la estructura mediante el comportamiento en el rango inelástico de sus miembros es sin duda la más importante en magnitud (ver figura 4.11), pues es la que ocurre cuando los elementos estructurales llegan al rango inelástico de su comportamiento desarrollando articulaciones plásticas, como se presenta en la figura 4.14.

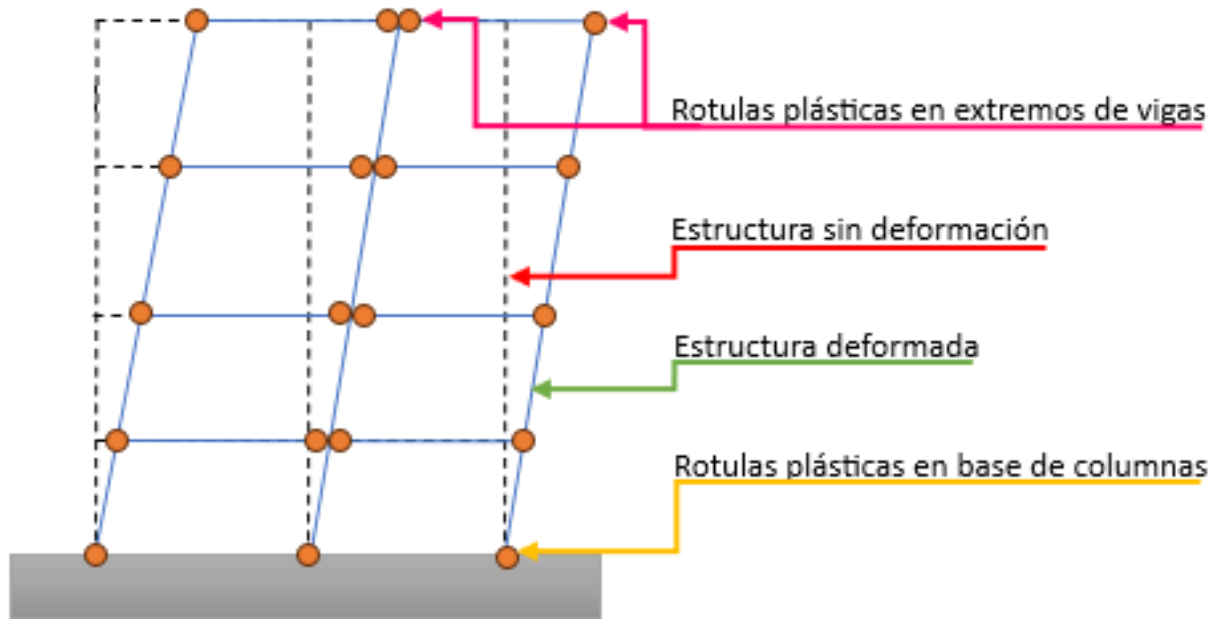


Figura 4.14: Deformación de la estructura.

El prototipo a utilizar representará un edificio de dos pisos, el cual será sometido a diferentes pruebas para obtener el comportamiento ante presencia de actividad sísmica, de modo que se obtendrá una base de datos que se utilizará para el aprendizaje y prueba de la red neuronal propuesta. Por lo cual se realizarán diez experimentos que se presentan en la tabla 4.2, donde el daño es causado por medio de pérdida de rigidez, a partir de aflojar una tuerca en el piso superior al cual se quiere dañar, de modo que se afectará la aceleración con la que se propaguen las ondas sísmicas.

Experimento	Piso Dañado	Datos obtenidos
1	Ninguno	Nominales
2	Ninguno	Nominales
3	Piso uno	Presencia de daño
4	Piso uno	Disminución de presencia de daño
5	Piso uno	Aumento de presencia de daño
5	Piso dos	Presencia de daño
6	Piso dos	Disminución de presencia de daño
7	Piso dos	Aumento de presencia de daño
8	Piso uno y dos	Presencia de daño en ambos piso
9	Piso uno y dos	Presencia de daño en ambos piso, aumento de daño en piso 1
10	Piso uno y dos	Presencia de daño en ambos piso, aumento de daño en piso 2

Tabla 4.2: Experimentos propuestos para el prototipo tipo edificio.

Las redes neuronales se presentan como una solución prometedora, debido a su capacidad para aprender y adaptarse a partir de ejemplos, les permite analizar grandes conjuntos de datos y extraer patrones significativos. Esto es especialmente útil en la detección de daño estructural, dado que los patrones detectados pueden ser sutiles y difíciles de identificar para los métodos tradicionales, dado que presentan limitaciones y su complejidad aumenta por el

tipo de estructura a analizar, lo que ha llevado a la necesidad de buscar utilizar este enfoque alternativo.

A lo largo de los años, este tipo de investigaciones y el desarrollo de esta tecnología, se han obtenido resultados impresionantes en comparación a métodos clásicos. Debido a su capacidad para mejorar la precisión del diagnóstico y reducir las posibilidades de falsos negativos o positivos, identificando así daños en estructuras que inicialmente podrían haber pasado desapercibidos.

Planteamiento para red neuronal monocapa

Las redes neuronales monocapa presentan ventajas como: la capacidad de trabajar con diferentes tipos de datos, ya sea información de sensores, imágenes o cualquier otra forma de entrada. Esto las hace versátiles y adaptables a una amplia gama de situaciones y aplicaciones. Con la continua investigación y mejora, han surgido las limitaciones en comparación a otros modelos de este tipo de algoritmos.

Por lo que, usando el principio de la "*navaja de Occam's*", conocido en inglés con el siguiente nombre; "*Occam's razor*", se explica que para resolver un problema: las soluciones más sencillas tienen mejores resultados, es decir, la red neuronal debe ser pequeña, fácil de entrenar y no ir más allá del problema que se desea resolver.

Por lo cual, es el resultado directo de la teoría básica de la probabilidad; dado que por definiciones básicas toda la suposición introduce posibilidades de error; pero si suposición no mejora la precisión de una teoría, su único efecto es aumentar la probabilidad de que la teoría general sea errónea.

En 1991, William H. Jefferys y James O. Berger generalizan y cuantifican el concepto de "suposiciones" de la formulación original, llegando a que una red neuronal con menos parámetros ajustables tendrá automáticamente una mejor probabilidad de mejora, debido al hecho de que las predicciones serán precisas, es decir, no tendrá que apartarse innecesariamente a posibles datos futuros que no se contemplen en la problemática inicial.

Teniendo así que diseñar las condiciones necesarias para poder comparar los resultados obtenidos por una red neuronal monocapa con algoritmos más complejos. En primer lugar, se requiere analizar la importancia y la complejidad del diagnóstico de daño estructural. Por lo que, es crucial detectar cualquier alteración en la integridad de la estructura, ya que puede tener significativas consecuencias para la seguridad y la vida útil de las mismas, afectando los resultados del estudio.

4.4.3. Diseño de red neuronal

Definición de la red neuronal

En esta etapa se selecciona el tipo de red neuronal que se utilizara con sus respectivos hiperparámetros, así como la recolección y selección de datos de entrada para la fase de aprendizaje y prueba.

Para llegar al tamaño óptimo en la red neuronal monocapa se usa la técnica de poda, es decir, se partió de una red neuronal con 60 neuronas y sus conexiones hasta llegar a los parámetros que se presentan en la tabla 4.3.

Característica	Selección
Tipo de red neuronal	Red neuronal monocapa
Número de neuronas en la capa de entrada	20
Función de activación	Función tangente hiperbólica
Grado de conectividad	Totalmente conectadas
Algoritmo de aprendizaje	Algoritmo de retropropagación
Índice de rendimiento	Error medio cuadrático (MSE)

Tabla 4.3: Parámetros finales de la red neuronal monocapa

Mientras que para la red neuronal multicapa se usa la técnica de crecimiento, es decir, se partió de los parámetros obtenidos con la red neuronal monocapa y se fueron agregando neuronas, conexiones y capas hasta llegar a los parámetros que se presentan en la tabla 4.4.

Característica	Selección
Tipo de red neuronal	Red neuronal multicapa
Número de capas ocultas	3
Número de neuronas en la capa de entrada	20
Número de neuronas en la capa oculta	18
Función de activación	Función tangente hiperbólica
Grado de conectividad	Totalmente conectadas
Algoritmo de aprendizaje	Algoritmo de retropropagación
Índice de rendimiento	Error medio cuadrático (MSE)

Tabla 4.4: Parámetros finales de la red neuronal multicapa.

Aprendizaje de la red neuronal

El aprendizaje de la red se realizará por medio del algoritmo de retropropagación. Por lo cual se requerirá una base de datos que contenga las frecuencias obtenidas por el prototipo con y sin daños ante excitación sísmica, para poder realizar alcanzar los pesos sinópticos óptimos para las neuronas y para las conexiones entre neurona.

Prueba de la red neuronal

En esta etapa se probará la red ante una base de datos que contenga diferentes frecuencias generadas por la presencia de daño en uno o los dos pisos del prototipo, siendo así una entrada diferente de datos a la utilizada durante el aprendizaje de la red. Donde la principal diferencia es que no se actualizan los pesos sinópticos de las neuronas como se realiza durante el procedimiento de aprendizaje.

El desempeño de la red se medirá a partir del índice de rendimiento de erro medio cuadrático (MSE), que determinara la eficiencia con la cual la red responde ante las entradas de excitación diferentes a la empleada durante su aprendizaje.

Mantenimiento de la red neuronal

Esta es la última fase en la metodología de la red neuronal, debido a que las redes normalmente son utilizadas para la solución de una sola problemática dinámica a la vez. Se

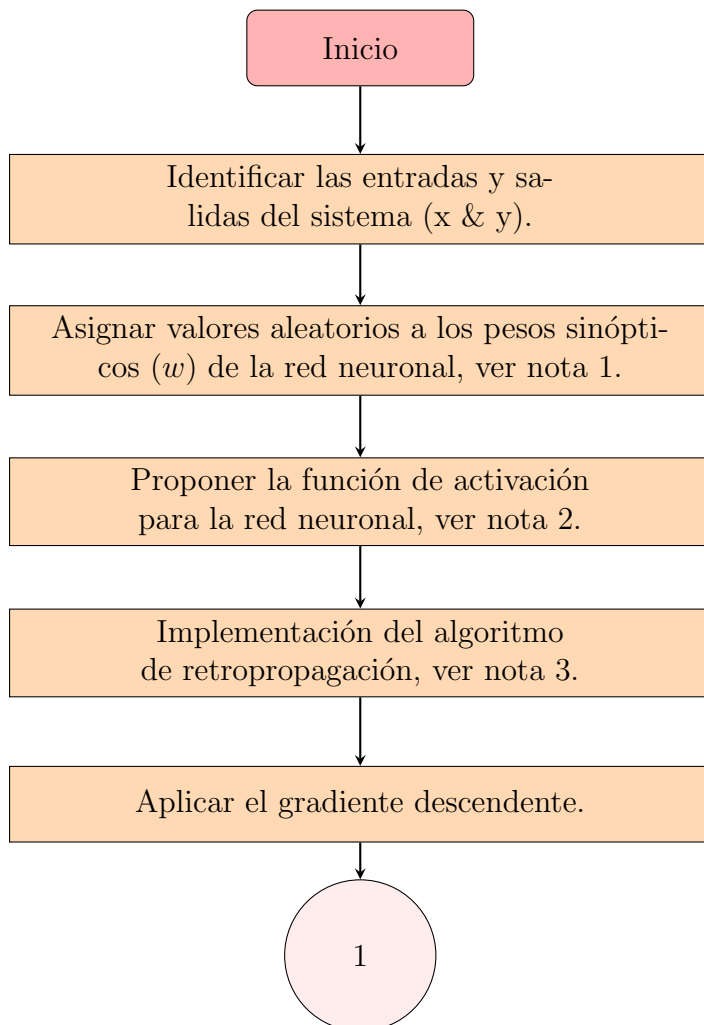
presenta la necesidad de validación continua a medida que pasa el tiempo, por lo cual, es indispensable que se valide la arquitectura utilizada con la finalidad de poder garantizar, ya que, al pasar el tiempo, surgirán nuevas bases de datos, las cuales pueden ser:

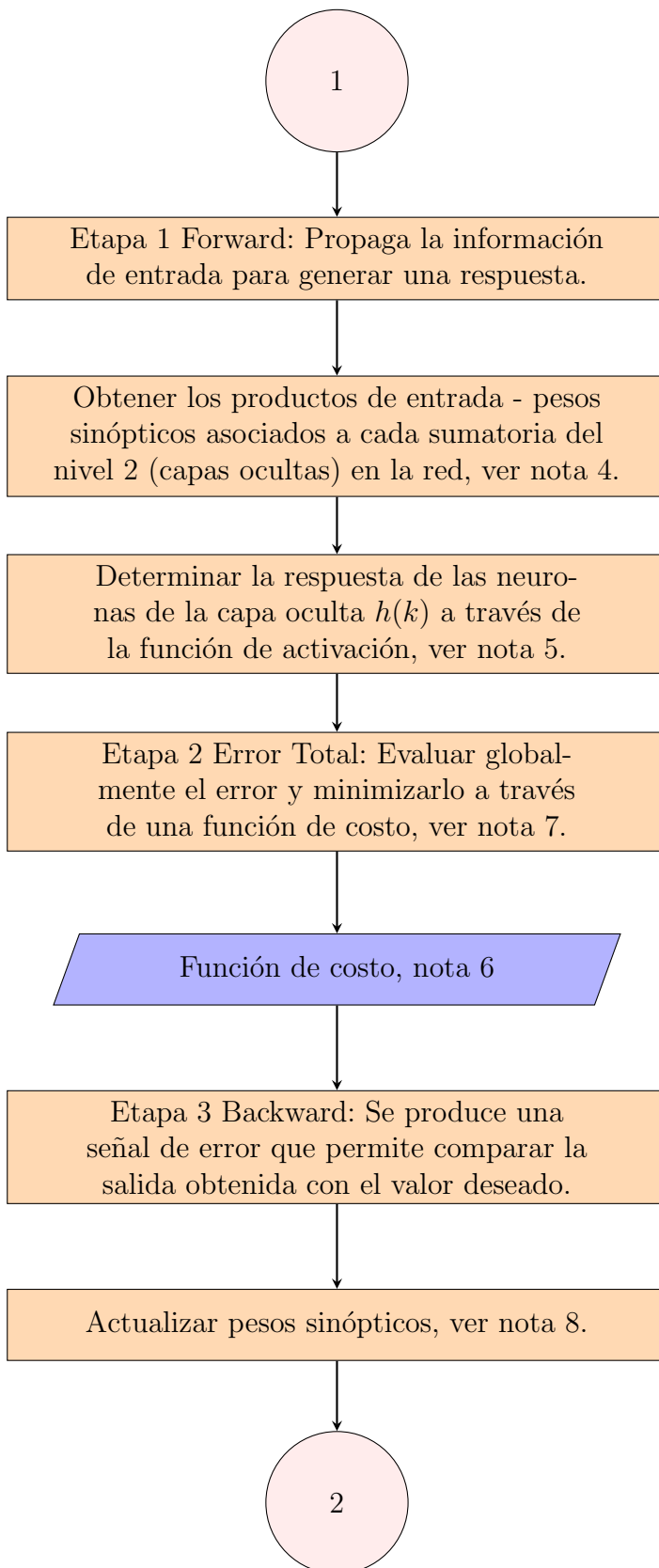
- Datos obtenidos a partir de experimentos o procesos reales.
- Modelos de simulación.
- Origen desconocido.
- A partir de instrumentos de medición, que proporcione un mejor margen de lectura que los anteriormente utilizados.

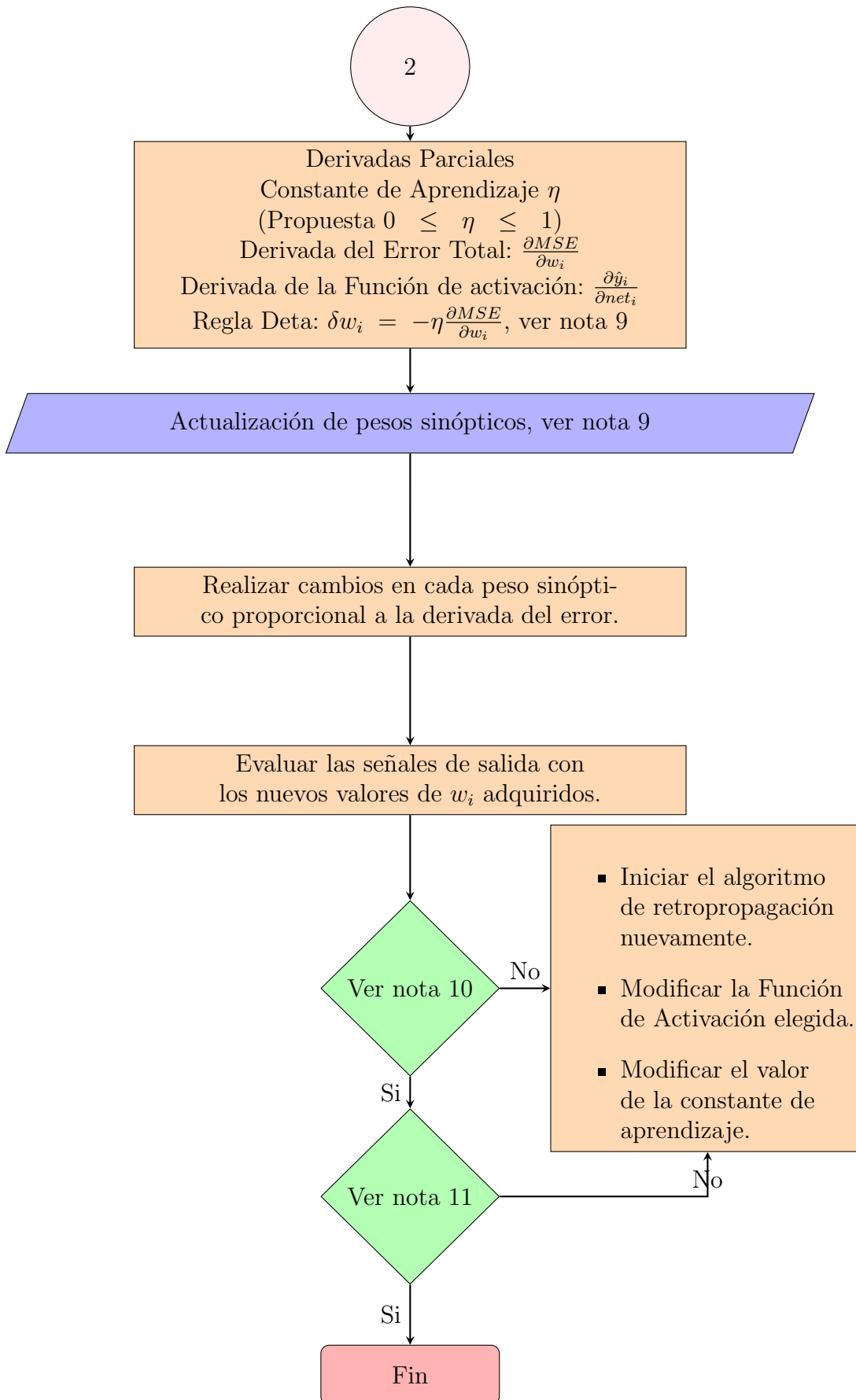
Presentado la oportunidad de realizar un nuevo proceso de aprendizaje, que permita mantener la red en un estado constante actualización para la mejora continua de esta herramienta de solución.

4.4.4. Fase de aprendizaje para redes neuronales

Se utilizará el aprendizaje por algoritmo de retropropagación, con la finalidad de minimizar la función de error por medio del gradiente descendiente. A continuación se presenta el diagrama de flujo en el que se describe el proceso de aprendizaje para las redes neuronales.







Notas

Nota 1. Condiciones iniciales $[-1,1]$.

Nota 2. Funciones de activación empleadas:

- Sigmoide
- Tangente hiperbólico

Nota 3. Algoritmo de retropropagación:

- Se refiere a la conexión entre neuronas de tipo propagación hacia delante o forward, como se presenta en la figura 4.15.

Trayectoria de izquierda a Derecha

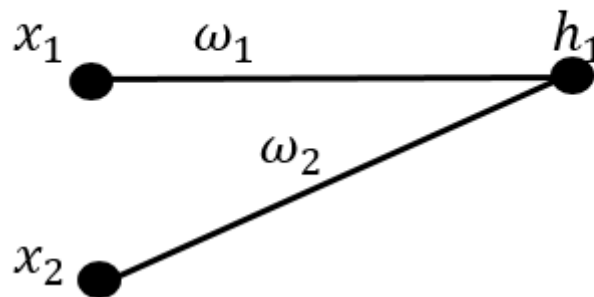


Figura 4.15: Trayectoria en la que se realiza la conexión entre neuronas de tipo propagación hacia delante.

- Error Total
- Se realiza la conexión entre neuronas de tipo propagación hacia atrás o backward.

Nota 4. Se emplea la regla de propagación.

Nota 5. Encontrar los valores estimados de las salidas de la red (Nivel 3).

Nota 6. Calcular la diferencia entre las salidas deseadas y las estimadas.

Nota 7. Función de costo, se considera como el error medio cuadrático (MSE), el cual es la sumatoria de errores parciales (depende del número de salidas).

Nota 8. La actualización de pesos sinópticos se realiza a partir de las derivadas parciales con la trayectoria que se presenta en la figura 4.16.

Trayectoria de Derecha a Izquierda

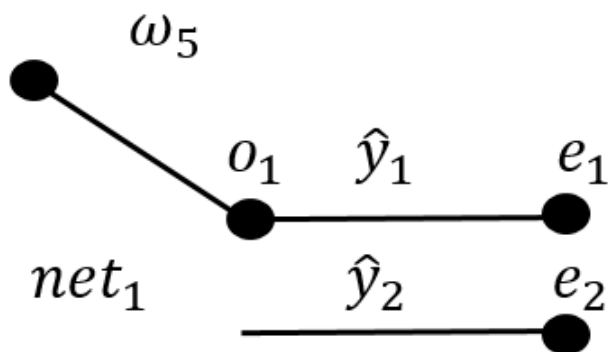


Figura 4.16: Trayectoria en la que se realiza la actualización de pesos sinópticos.

Nota 9. Consideración de signo, nos determina el que signo que debe presentar el error dependiendo de la comparación que se realiza entre la salida estima y la salida desea:

$$\begin{aligned} \hat{y}_i - y_i = + &\rightarrow \text{el error debe ser mayor a } 0 \\ y_i - \hat{y}_i = - &\rightarrow \text{el error debe ser menor a } 0 \end{aligned} \tag{4.47}$$

Nota 10. ¿El margen de error se ha minimizado?

Nota 11. ¿Las salidas estimadas son iguales a las deseadas?

Nota 12. El número de interacciones está sujeto al criterio del diseñador.

4.5. Diagnóstico de daño con redes neuronales

4.5.1. Red neuronal monocapa

Resultados

En este caso, al utilizar una red monocapa, se requiere implementar *Occa's razor*, para así obtener resultados que llegarían a compararse con estructuras de redes neuronales más complejas que presenten mejores resultados, por esta razón se necesitó una adecuada selección de los datos de entrenamiento, así como de los hiperparámetros de la red, como se presenta en la tabla 4.5. Donde se fijó una taza de aprendizaje, así como la función de activación, para posteriormente ir reduciendo el número de neuronas hasta llegar a los parámetros que se presentaron en la tabla 4.3.

Piso Dañado	α	Número de neuronas	Función de activación	Problema que presento
2	0.1	60	Sigmoide	Durante el aprendizaje entregó MSE menores al 3×10^{-4} , pero al modificar la gravedad del daño se presentó sobreaprendizaje, provocando un MSE en el piso 2 de $6,8426 \times 10^{-1}$.
2	0.1	40	Sigmoide	Durante el aprendizaje entregó MSE menores al 3×10^{-2} , pero al disminuir la gravedad del daño se presentó un MSE en el piso 2 de $4,7426 \times 10^{-1}$.
2	0.1	20	Sigmoide	Durante el aprendizaje entregó MSE menores al 3×10^{-1} , pero al disminuir y aumentar la gravedad del daño se presentó un MSE en el piso 2 de $4,7452 \times 10^{-1}$ y $3,4183 \times 10^{-1}$ respectivamente.
2	0.1	60	Tangente hiperbólico	Durante el aprendizaje entregó MSE menores al 3×10^{-12} , pero al disminuir la gravedad del daño se presentó un MSE en el piso 4,627.
2	0.1	40	Tangente hiperbólico	Durante el aprendizaje entregó MSE menores al 3×10^{-10} , pero al aumentar la gravedad del daño se presentó un MSE en el piso 0,0671.
2	0.1	20	Tangente hiperbólico	Durante el aprendizaje entregó MSE menores al 3×10^{-9} , pero al disminuir y aumentar la gravedad del daño se presentó un MSE en el piso 2 de $4,5816 \times 10^{-9}$ y $3,4183 \times 10^{-9}$ respectivamente.

Tabla 4.5: Ajuste de parámetros red monocapa.

Se presenta en la figura 4.17 la respuesta de la red neuronal monocapa cuando se emplea el ajuste de parámetros resaltados en color azul de la tabla 4.5. En este caso se tiene que la fuerza de histéresis estimada es mayor a la referencia, esto ocasiona que al momento de generarse los ciclos de histéresis los daños diagnosticados sean más severos de lo que realmente son.

Debido al fijar la función de activación en sigmoide he ir disminuyendo el número de neuronas, se tiene que el MSE en el aprendizaje aumenta, pero al cambiar la gravedad del daño se registra que el MSE se mantiene cerca del error de aprendizaje lo que significaría que la sensibilidad de la red neuronal ante diferentes entradas es buena, pero a mayor número de neuronas el error aumenta debido a que la sensibilidad de la red disminuye.

Lo que significaría que al momento de probarlo en otros escenarios de daño se presentarían falsos positivos debido a que agrava las condiciones en las que se encuentra la estructura. Lo

que significaría pérdida de tiempo y recursos debido a que se requería mantenimiento en los daños falsamente diagnosticados.

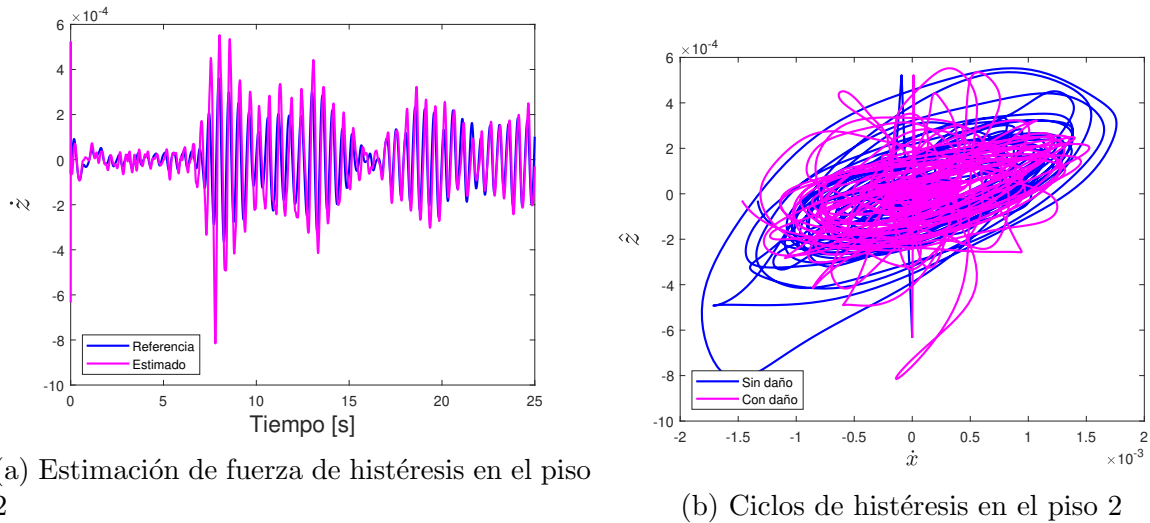


Figura 4.17: Resultados obtenidos al utilizar red monocapa con función de activación sigmoide.

Mientras que en la tabla 4.6, se presenta el desempeño de la red neuronal monocapa de acuerdo al ajuste de parámetros resaltados en color rosa de la tabla 4.5

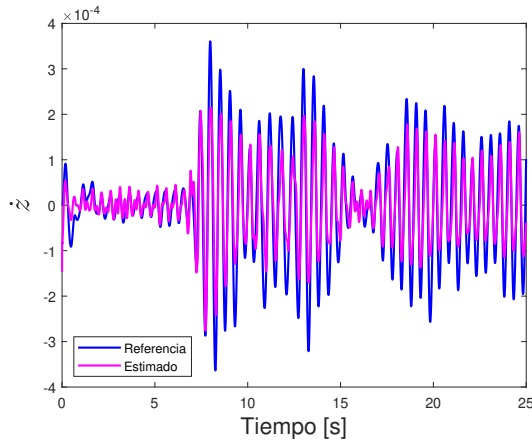
Piso Dañado	α	Tiempo de aprendizaje [s]	MSE piso 1	MSE piso 2
2	0.1	61.7350	$1,00859 \times 10^{-8}$	$3,3394 \times 10^{-9}$
2	0.001	61.728661	$1,0089 \times 10^{-8}$	$3,3394 \times 10^{-9}$
1 y 2	0.1	70.1482	$1,0085 \times 10^{-8}$	3.8129
1 y 2	0.001	78.8224	$1,0085 \times 10^{-8}$	3.8130

Tabla 4.6: Resultados obtenidos al utilizar red monocapa.

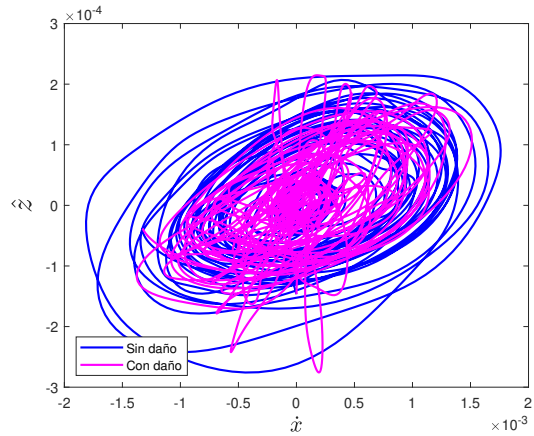
En la figura 4.18 se presenta la respuesta obtenida cuando el piso dos se encuentra dañado, donde el índice de rendimiento obtenido permite considerar esta configuración de red neuronal como una opción en el diagnóstico de daño estructural. Sin embargo, al tener daño estructural en ambos pisos (ver figura 4.19) se requiere un aprendizaje más riguroso debido a que el MSE aumenta en el piso dos.

Esto se debe a cómo se disipa la energía en una estructura, ya que cuando presenta una excitación sísmica de alta intensidad, esta afecta más a las estructuras que cuentan con un mayor número de pisos. Debido a que el centro de gravedad se encontrará en movimiento, lo que provoca que la estructura pueda vibrar a su frecuencia de resonancia.

Mientras que a excitaciones sísmicas de baja intensidad afectan más a las estructuras que tiene pocos pisos. Ya que las estructuras deben poseer una gran ductilidad para poder incurrir en grandes deformaciones dentro del rango inelástico para formar articulaciones plásticas en sitios estratégicos.

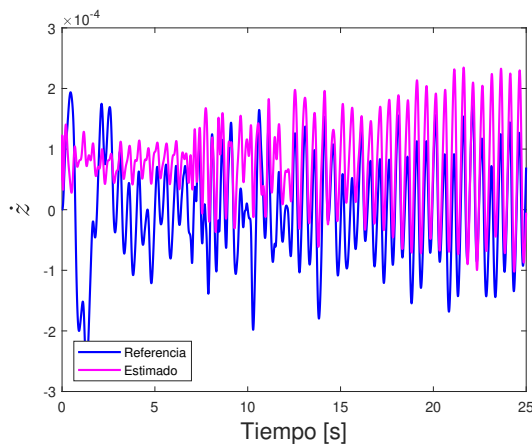


(a) Estimación de fuerza de histéresis en el piso 2

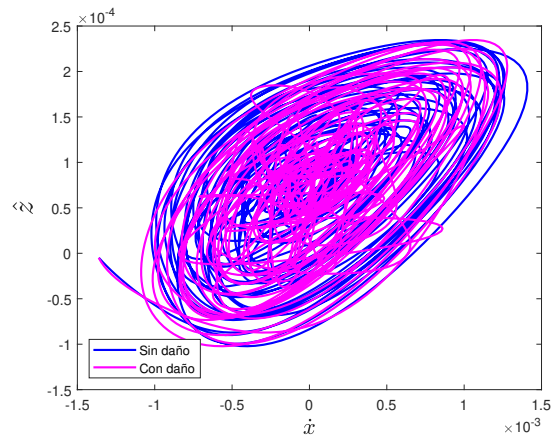


(b) Ciclos de histéresis en el piso 2

Figura 4.18: Resultados obtenidos al utilizar red monocapa con daño en el piso 2.



(a) Estimación de fuerza de histéresis en el piso 1



(b) Ciclos de histéresis en el piso 1

Figura 4.19: Resultados obtenidos al utilizar red monocapa con daño en el piso 1 y piso 2.

Por lo que implementar redes neuronales monocapa en estructuras con mayor número de pisos significaría la recolección de una base de datos que permita el correcto aprendizaje de la red, aumentando el peso computacional que se requiere para realizar el análisis de los datos de forma correcta, teniendo en mente el no caer en el sobreaprendizaje al punto que solo se pueda identificar el daño estructural ante una fuerza de excitación externa específica por lo cual se volvería inservible ante otros casos de estudio.

Análisis

En conclusión, las redes neuronales monocapa, a pesar de sus limitaciones como lo son: sobreaprendizaje, selección de hiperparámetros, peso computacional, sensibles a la escala de los datos y un mayor tiempo de aprendizaje para lograr los índices de rendimiento presentados, por mencionar algunas. Así mismo su eficacia para detectar el de daño estructural a partir de alteraciones en el comportamiento que presenta una estructura ante la presencia de

activada sísmica, a mejorando la precisión y reduciendo de los falsos positivos y negativos. Aunque con sus limitaciones, su versatilidad y capacidad para analizar diferentes tipos de datos las convierte en una herramienta invaluable en el campo de la ingeniería estructural. En definitiva, su uso promete seguir creciendo y contribuyendo al desarrollo de estructuras más seguras y duraderas.

Es importante destacar que, si bien las redes neuronales monocapa presentan varias ventajas ante modelos tradicionales, también tienen limitaciones. Ya que, su rendimiento depende en gran medida de la fiabilidad y la cantidad de datos de aprendizaje disponibles. Además, la interpretación de los resultados puede ser un desafío, ya que las redes neuronales a menudo se consideran “cajas negras”. Esto significa que pueden ser difíciles de entender y explicar, lo que puede generar desconfianza en algunos sectores.

4.5.2. Evaluación con red neuronal multicapa

Resultados

Al utilizar una red multicapa, se requiere una adecuada selección hiperparámetros de la red, como se presenta en la tabla 4.7. Donde se tomó como referencia el ajuste de parámetros que entregó mejor índice de rendimiento en la red neuronal monocapa, para posteriormente ir aumentando el número de capas hasta llegar a los parámetros que se presentaron en la tabla 4.4.

Piso Dañado	α	Número de capas ocultas	Función de activación	Problema que presento
2	0.1	1	Tangente hiperbólico	Durante el aprendizaje entregó MSE menores al 3×10^{-7} , pero al disminuir la gravedad del daño se presentó un MSE en el piso $3,3744 \times 10^{-5}$.
2	0.1	2	Tangente hiperbólico	Durante el aprendizaje entregó MSE menores al 3×10^{-9} , pero al aumentar la gravedad del daño se presentó un MSE en el piso $3,2471 \times 10^{-7}$.
2	0.1	3	Tangente hiperbólico	Durante el aprendizaje entregó MSE menores al 3×10^{-9} , pero al disminuir y aumentar la gravedad del daño se presentó un MSE en el piso 2 de $3,5816 \times 10^{-9}$ y $3,4183 \times 10^{-9}$ respectivamente.

Tabla 4.7: Ajuste de parámetros red neuronal multicapa.

Preliminarmente, se puede llegar que al estimar la fuerza de histéresis, con el método propuesto es prometedor debido a que permite utilizar mediciones disponibles y no se requiere la necesidad de instrumentación adicional.

Dado que se pueden estimar los parámetros como: posición y velocidad para calcular la energía disipada por la estructura ante excitación sísmica. Y así diagnosticar oportunamente

los daños que se puedan generar sin la necesidad de requerir pruebas externas que verifiquen su integridad estructural.

Como se presenta en la tabla 4.7, al mantener los parámetros fijos previamente obtenidos en la red neuronal monocapa e ir incrementando capas ocultas, se presenta que el MSE durante el aprendizaje es menor al 3×10^{-7} y 3×10^{-9} , para una y dos capas ocultas, pero al modificar la gravedad del daño se aumenta el error. En el caso de tener tres capas ocultas, el error obtenido en la fase de aprendizaje y al momento de modificar la intensidad del daño se mantiene en el mismo rango de error, por lo cual esta será el ajuste de parámetros a utilizar. Mientras que en la tabla 4.8, se presenta el desempeño de la red neuronal multicapa de acuerdo al ajuste de parámetros resaltados en color verde de la tabla 4.7

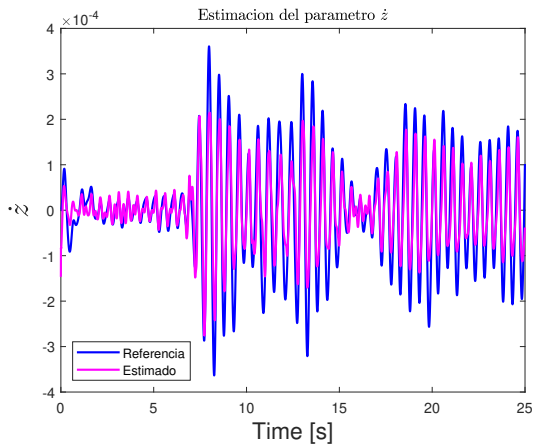
Piso Dañado	α	Tiempo de aprendizaje [s]	MSE piso 1	MSE piso 2
2	0.1	32.2138	$1,0085 \times 10^{-8}$	$3,3394 \times 10^{-9}$
2	0.001	33.126263	$1,0085 \times 10^{-8}$	$3,3394 \times 10^{-9}$
1 y 2	0.1	52.5646	$1,008598 \times 10^{-8}$	$3,3394 \times 10^{-9}$
1 y 2	0.001	43.2457	$1,00859 \times 10^{-8}$	$3,3394 \times 10^{-9}$

Tabla 4.8: Resultados obtenidos al utilizar red neuronal multicapa.

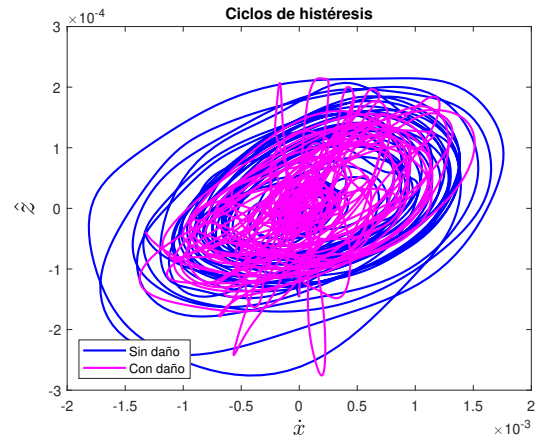
A partir de los resultados obtenidos en la tabla 4.8, se tiene que al dañar uno de los dos pisos del prototipo, el MSE en ambos pisos se encuentra en el rango de $1 - 3 \times 10^{-9}$. Lo que significa la red neuronal está entregando un buen desempeño tanto en la etapa de aprendizaje como de prueba, ya que aún cuando se modifique la gravedad del daño, el único factor que presenta un aumento es en tiempo de aprendizaje que se requiere lo que significa que la red neuronal multicapa tiene la sensibilidad adecuada para adecuarse a la modificación de datos de entrada.

Mientras que al encontrarse dañados los dos pisos del prototipo, en el caso de aumentar el daño en el piso uno, como se presenta en la tabla 4.8 resaltado en color amarillo, se tiene que el tiempo de aprendizaje se aumenta. Pero este efecto se puede contrarrestar al modificar la constante de aprendizaje, como se presenta en la tabla 4.8 en color naranja.

Lo que significa que la red neuronal multicapa permite modificar el tiempo que se requiere para realizar el diagnóstico sin afectar su precisión y eficiencia, reduciendo así los falsos positivos y negativos.

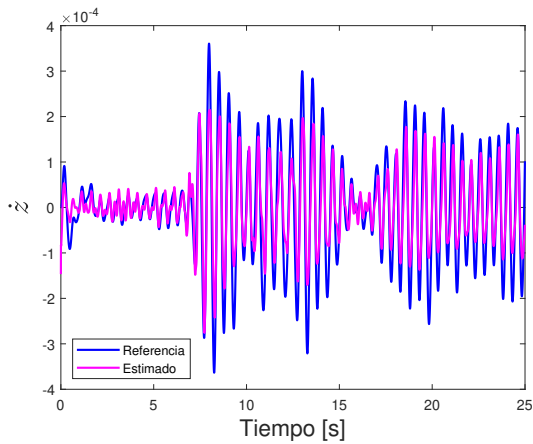


(a) Estimación de fuerza de histéresis en el piso 2

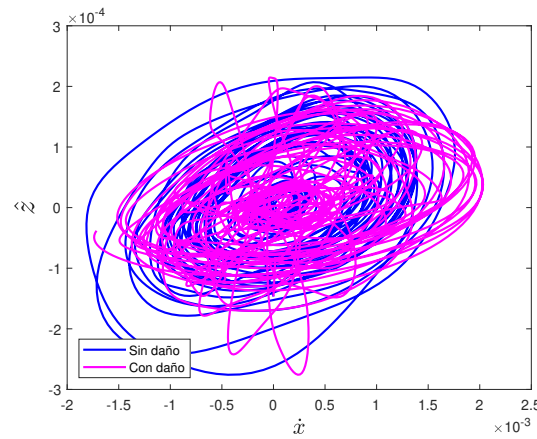


(b) Ciclos de histéresis en el piso 2

Figura 4.20: Resultados obtenidos al utilizar red neuronal multicapa con daño en el piso 2.



(a) Estimación de fuerza de histéresis en el piso 2



(b) Ciclos de histéresis en el piso 2

Figura 4.21: Resultados obtenidos al utilizar red neuronal multicapa con daño en el piso 1 y piso 2.

Análisis

La explicación de los ciclos de histéresis en edificios con daño estructural se puede basar en el análisis de la curva de histéresis obtenida durante pruebas experimentales o simulaciones computacionales. Esta curva muestra la relación entre la fuerza aplicada a la estructura y la respuesta correspondiente en términos de desplazamiento. En un edificio sin daño estructural, se espera que la curva de histéresis sea simétrica y que la energía disipada durante el ciclo de carga y descarga sea mínima.

Sin embargo, en un edificio con daño estructural, la curva de histéresis puede presentar asimetrías y menor energía disipada (ver figura 4.13). Estas asimetrías son indicativas del daño existente en la estructura y reflejan la pérdida de rigidez y resistencia de los elementos estructurales afectados.

Aunque el comportamiento de los ciclos de histéresis en edificios con daño estructural puede

variar dependiendo del tipo y grado de daño presente. Por ejemplo, para una estructura con fisuras en los muros, los ciclos de histéresis pueden mostrar un aumento en la rigidez inicial, seguido de una rápida pérdida de resistencia. Por otro lado, en una estructura con daño en los pilares, los ciclos de histéresis pueden mostrar una disminución gradual en la rigidez y resistencia durante los ciclos de carga y descarga, como se presentan en las figuras 4.18 - 4.21.

Por lo tanto, este enfoque en los ciclos de histéresis en edificios con daño estructural proporcionan una explicación detallada sobre el comportamiento de las estructuras ante cargas sísmicas. En consecuencia, el estudio de los ciclos de histéresis en edificios con daño estructural es esencial para promover la seguridad y la resiliencia en las infraestructuras ante sismos. En conclusión, el análisis de estas curvas es fundamental para evaluar la capacidad de resistencia y la integridad estructural de los edificios afectados por actividad sísmicas. Por lo que al entender los ciclos de histéresis permite a los ingenieros identificar el tipo y grado de daño existente, así como diseñar estrategias de rehabilitación y refuerzo adecuadas.

Capítulo 5

Conclusiones y trabajos futuros

5.1. Conclusiones

Las redes neuronales multicapa son alternativa efectiva para el diagnóstico de daño en estructuras debido a su capacidad para aprender patrones complejos y no lineales en los datos de entrada. Al contrario de usar una red neuronal monocapa en la que se debe de diseñar el experimento, de tal manera que se llegue a tener índices de rendimiento como el error medio cuadrático (MSE) mejores o parecidos a una red neuronal multicapa.

Llegando a que, si no se presentan las mismas condiciones, en las que se diseñó el experimento, la red neuronal monocapa entrega resultados medianamente satisfactorios, como pasó en el caso de presentar daño en los dos pisos del prototipo tipo edificio.

Mientras que al utilizar, una red neuronal multicapa, se presentan una serie de ventajas adicionales en el campo del diagnóstico de daño en estructuras, ya que, son capaces de lidiar con grandes cantidades de datos, permitiendo un mejor aprendizaje de la red ante la presencia de diversos casos de daño.

Por lo que, este tipo de redes neuronales son las más adecuadas para el análisis y procesamiento de la información recopilada por el sistema de monitoreo estructural, entregando así, mejores índices de rendimiento como lo es el error medio cuadrático (MSE) sin importar cuál sea el piso del prototipo que presente daño o si ambos piso presenten daño. Debido a su robustez para resistir los errores y/o perturbaciones en los datos de entrada, permite poder realizar el monitoreo en línea, fuera de línea o una combinación de aprendizaje fuera de línea y validación en línea, como fue el caso de este trabajo.

Finalmente, es importante destacar que el modelo de Bouc-wen, presenta una complejidad computacional. Debido a la naturaleza no lineal del modelo, los cálculos requeridos para su implementación pueden volverse computacionalmente intensivos, lo que puede limitar su en tiempo real por tiempos prolongados. Además, de la necesidad de ajustar varios parámetros del modelo, como la rigidez y el coeficiente de histéresis, dificultando su implementación y requiriendo un conocimiento especializado para realizar correctamente los ajustes.

5.2. Trabajo futuro

Considerando los resultados obtenidos, al utilizar redes neuronales multicapa en este trabajo, se puede considerar métodos metaheurísticos que aún no se han aplicado para identificación de sistemas o implementar otras metodologías de computación inteligente para la caracterización del amortiguador, así como seleccionar un mejor método de procesamiento para los datos debido a la gran cantidad que se requiere analizar.

Otros trabajos futuros se contempla abordar la modificación del paradigma de aprendizaje y la simplificación del algoritmo implementado. Para su posible implementación en otras áreas de estudio o siguiendo esta línea de estudio, implementarlo en modelos estructurales con más de dos grados de libertad, con la finalidad de poder implementar en edificios de la ciudad y así conseguir un modelo más realista que pueda detectar con mayor eficiencia los daños causados por actividad sísmica.

Por último, es importante considerar las desventajas asociadas con el uso una red neuronal multicapa, como la necesidad de conocimientos especializados en la problemática a tratar. Los trabajos futuros deben centrarse en mejorar la eficacia, así como en explorar su integración con otros métodos de histéresis enfocados en diagnóstico.

Apéndice A

Apéndice A: Características del prototipo tipo edificio

El prototipo tipo edificio de dos pisos utilizado en este estudio, fue fabricado en aluminio, tiene unas dimensiones (32,5 x 53) cm y una altura de 1,2 m. La sección transversal de las columnas es rectangular con un ancho de (0.635×2.54) cm. En la tabla A.1 se presenta un desglose de los componentes utilizados para la obtención de la base de datos para el aprendizaje y prueba del método de diagnóstico de daño.

Característica	Descripción
Altura del primer piso	58 cm
Altura del segundo piso	62 cm
Duración de excitación sísmica	25 s
Servomotores Quanser	Modelo I-40
Acelerómetros Analog Dvice	Modelo XL403A 1-15 g [1x800]Hz
Tarjetas Electronicas Intenco	DAC/USB2
Programas de Operación	Windows 7 Matlab 2012b/Simulink
k_1	12011N/m
k_2	12108N/m
m_1	2,034kg
m_2	2,534kg
f_i	1.758Hz
f_2	4.0Hz
x_i	$i = 1, 2, \dots, n$

Tabla A.1: Características del prototipo tipo edificio de 2 pisos.

En la tabla A.2, se presenta el escalamiento realizado para la adquisición de datos en el prototipo, ya que si no se realiza este procedimiento, los resultados obtenidos estarían comprometidos debido a que una estructura de las dimensiones propuestas no reaccionaría de igual forma ante una fuerza sísmica como la registrada.

	Movimiento sísmico original	Movimiento sísmico escalado
Desplazamiento	16.92	1.50
Duración	39.98	11.91
Tamaño de muestreo	2167	2167

Tabla A.2: Escalamiento de movimiento sísmico Northridge para el prototipo tipo estructura.

Finalmente, debido ante la presencia de actividad sísmica, solo se cuenta con instrumentos de medición que proporcionan la aceleración, se requiere estimar los parámetros de velocidad y el desplazamiento. Por lo cual, son estimados empleando un filtro pasa altas, para obtener las señales restantes, donde las características específicas son: frecuencia de corte que se establece en 0,3 Hz para eliminar los componentes de baja frecuencia y evitar la derivada, obteniendo así:

$$f(s) = \frac{s^2}{s^2 + 3,77 + 3,55} \times \frac{s^2}{s^2 + 3,77 + 355} \times \frac{1}{s} \quad (\text{A.1})$$

Bibliografía

- [1] Rubén Boroschek K., Tomás Núñez A., Rodrigo Carreño V., and Alfonso Larraín V. Investigación edificios emblemáticos diagnóstico e identificación de daño estructural. *BIT*, 73, 2010.
- [2] CENAPRED. *Evaluación de la seguridad estructural de edificios*, pages 11–200. Coordinación Nacional de protección civil México, 24 de Junio 2013.
- [3] Servicio Sismológico Nacional Instituto de Geofísica UNAM. Zona de subducción mexicana y su potencial para un sismo mayor www.ssn.unam.mx/sismicidad/reportes-especiales/otros/SSNRE-Subduccion-y-potencial-de-sismo-mayor.pdf.
- [4] Servicio Geológico Mexicano. Sismología de México www.sgm.gob.mx, 22 de Marzo de 2017.
- [5] Certicalia. ¿qué es el diagnóstico estructural? www.certicalia.com/diagnostico-estructural-que-es-el-diagnostico-estructural, 29 de Noviembre de 2022.
- [6] Ricardo Andrés Saavedra Guzmán. Análisis de algunas causas de fallas y defectos de las estructuras. Master's thesis, Escuela Colombiana de Ingeniería Julio Garavito, 6 de Agosto de 2019.
- [7] Dr. Arquitecto Enrique Maya Mirandá. Rehabilitación estructural de conjuntos históricos. *Universidad de Navarra*, 8 de Febrero de 2019.
- [8] Diego Miramontes de León, Hipólito Gurrola Carrillo, and Brenda Araceli Castro de la Torre. Diagnóstico estructural en monumentos y edificios históricos en la ciudad de zacatecas. *XVI Congreso Nacional de Ingeniería Estructural*, 2008.
- [9] Roberto Meli. *Diagnóstico y Rehabilitación Estructural de Edificios Históricos*, pages 7–11. Instituto de Ingeniería UNAM, 2016.
- [10] Farrar C. R. and Doebling S. W. and Nix D. A. Vibration-based structural damage identification. *Philosophical Transaction Of The Royal Society*, 359:131–149, 2001.
- [11] Doebling S. W. and Farrar C. R. and Prime M. B. A summary review of vibration-based damage identification methods. shock and vibration digest. *The Shock and Vibration Digest*, 30:91–105, March 1998.
- [12] Doebling S. W., Farrar C. R., Prime M. B., and Shevitz D. W. Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics. *Los Alamos National Laboratory*, 1996.

- [13] Farrar C. R. and Worden K. An introduction to structural health monitoring. *Philosophical Transaction Of The Royal Society*, 365:303–315, 2007.
- [14] Alberto Renán Ancona Lazcano, Rolando Salgado Estrada, Sergio Aurelio Zamora Castro, and Fernando Marcial Martínez. Evaluación de métodos de detección de daño en estructuras mediante el uso de vibraciones. *Congreso Nacional de Ingeniería Sísmica*, XVIII, 2011.
- [15] Pablo Andrés Rojas Jaramillo. Análisis modal del banco de ensayo de vibraciones del laboratorio del departamento de ingeniería mecánica. Master's thesis, Universidad del Bío-Bío Facultad de Ingeniería Departametro de Ingeniería Mecánica, 2014.
- [16] Hernández Hugo, Huergo Iván, and Guerrero Luis. Formas modales de edificios altos obtenidas por medio de modelos continuos. *XXII Congreso Nacional de Ingeniería Sísmica*, 2019.
- [17] David Henao Ángel, Juan Carlos Botero Palacio, and David Muriá Vila. Identification of dynamic properties of a structural model subjected to ambient vibration and forced vibration using shaking table. *Sociedad Mexicana de Ingeniería Sísmica A.C.*, 91:54–73, 2014.
- [18] Das S. Saha P. and Patro S. Vibration-based damage detection techniques used for health monitoring of structures: a review. *Journal of Civil Structural Health Monitoring*, 6:477–507, 2016.
- [19] Farrar C., Doebling S. W. and Cornwell P., and Straser E. G. Variability of model parameter measured on the alamosa canyon bridge. *Proceedings of SPIE- The International Society for Optical Engineering*, XV International Modal Analysis Conference:1–34, 1997.
- [20] Doebling S. W. and Farrar C. R. and Goodman R. S. Effects of measurement statics on the detection of damage in the alamosa canyon bridge. *XV International Modal Analysis Conference*, pages 912–929, 1997.
- [21] Sohn H. and Farrar C. R. Damage diagnosis using time series analysis of vibration signals. *Smart Mater. Struct.*, 10:1–6, 2001.
- [22] L. M. Villa García. Comparison of the estimated modal parameters of structures from classical and operational modal analysis with mass changes. *Revista Internacional de Métodos Numéricos para Cálculo y Diseño en Ingeniería*, 33:188–196, 2017.
- [23] Clinton J. F., Bradford S. C., Heaton T. H., and Favela J. The observed wander of the natural frequencies in a structure. *Bulletin of the Seismological Society of America*, 90:237–257, February 2006.
- [24] Doyle J.F. *Wave Propagation in Structures*. Springer-Verlag, New York, 2 edition, 1997.
- [25] Jesús Morales Valdez, Luis Alvarez Icaza, and Antonio Concha. Identificación paramétrica en tiempo real de edificios mediante propagación de ondas. *Memorias del Congreso Nacional de Control Automático*, pages 28–30, 2016.

- [26] Jesús Morales Valdez, Luis Alvarez Icaza, and Antonio Concha. Identificación de rigidez en los entresijos de un edificio mediante tiempos de viaje de onda. *Memorias del Congreso Nacional de Control Automático*, XVI:14–17, 2014.
- [27] Zapico J., González M., and k. Worden. Damage assesment using neural networks. *Mechanical Systems and Signal Processing*, 17:119–125, 2003.
- [28] Jesús D. Villalba, Ivan D. Gómez, and José E. Laier. Damage detection in beams by using artificial neural networks and dynamical parameters. *Facultad de Ingeniería, Universidad de Antioquia*, 63:141–153, 2012.
- [29] Jesús Morales Valdez, Mario Lopez Pacheco, and Wen Yu. Automated damage location for building structures using the hysteretic model and frequency domain neural networks. *Structural Control and Health Monitoring*, 27, 2019.
- [30] R.Hecht-Nielsen. *Neurocomputing: Picking the Human Brain*. IEEE Spectrum, Marzo 1988.
- [31] T Kohonen. *An Introduction to Neural Computing*. Neural Networks, Vol.1, 1988.
- [32] James A. Freeman and David M. Skapura; versión en español de Rafael García Bermejo Giner; con la colaboración de Luis Joyanes Aguilar. *Redes neuronales: algoritmos, aplicaciones y técnicas de programación*. Addison-Wesley Iberoamericana, 1993.
- [33] Ana Barragán Montero. ¿qué es una red neuronal artificial? https://cebebelgica.es/es_ES/blog/10/que-es-una-red-neuronal-artificial.html, 2010.
- [34] Bonifacio Martín del Brío and Alfredo Sanz Molina. *Redes Neuronales y Sistemas Borrosos*. RA-MA, 2001.
- [35] Ligdi González. ¿qué es el perceptrón? perceptrón simple y multicapa <https://aprendeia.com/que-es-el-perceptron-simple-y-multicapa/>, 2023.
- [36] Sergio Martínez Hamdoun. Estrategia evolutiva para diseñar redes neuronales convolucionales, Junio 2021.
- [37] Krizhevsky Alex, Sutskever Ilya, and Hinton Geoffrey. Imagenet classification with deep convolutional neural networks. *Neural Information Processing Systems*, 25, 01 2012.
- [38] Diego calvo. Clasificación de redes neuronales artificiales <https://www.diegocalvo.es/clasificacion-de-redes-neuronales-artificiales/>, 2017.
- [39] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *ICLR*, 2015.
- [40] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. *arXiv*, 2014.
- [41] Jesús Pérez Guerrero. *Redes recurrentes*, 2020.

- [42] Mariano Rivera. Introducción a redes neuronales recurrentes (rnn) perceptrón simple y multicapa http://personal.cimat.mx:8181/~rivera/cursos/aprendizaje_profundo/RNN_LSTM/introduccion_rnn.html#red-de-memoria-corta-a-largo-plazo-long-term-short-memory-lstm, 2018.
- [43] Sepp Hochreite and Jurgen Schmidhuber. Long short-term memory. *Neural Computation* 9(8):1735-1780, 1997.
- [44] Kyunghyun Cho, Bart van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014.
- [45] Simon Haykin. *Neural Networks and Learning Machines*. Pearson Prentice Hall, third edition, 1999.
- [46] François Chollet. *Deep Learning with Python*. Manning Publications Co, 2018.
- [47] Xabier Basogain Olabe. *Redes neuronales artificiales y sus aplicaciones*.
- [48] José Ramón Hilera González and Victor José Martínez Hernando. *Redes neuronales artificiales: fundamentos, modelos y aplicaciones*. RA-MA Editorial, January 1995.
- [49] Richard Lacks, Dagmar Mack, Juergen Ziola, and Keith Ahern. *Neuronal Networks: Basics and Applications*. Springer Electronic Media, 1998.
- [50] Pablo L. Ringegni Andrés and Martínez del Pezzo. Mecánica y mecanismos: Vibraciones de un grado de libertad. Technical report, Universidad Nacional de La Plata, Departamento de Aeronáutica, Facultad de Ingeniería, 2018.
- [51] Paul E. Tippens. *Física: Conceptos y Aplicaciones*. McGRAW-HILL, 2011 7ma Edición.
- [52] M.C. Fernando Javier Elizondo Garza. Práctica número 3: Frecuencias naturales. Laboratorio Vibraciones mecánicas de Universidad Autónoma de Nuevo León, Facultad de Ingeniería Mecánica y Eléctrica.
- [53] Snill K. Chopra. *Dynamics of structures: theory and applications to earthquake engineering*. Pearson Prentice Hall, 2017.
- [54] Antonio Pérez. Amortiguamiento. Digital, 2006.
- [55] Laboratorio Virtual de Riesgos Laborales Universidad de Córdoba. Equipos de medida de vibraciones www.uco.es/RiesgosLaborales/fisicoyquimico/vibraciones/tutorials/view/6-Equipos-de-medida-de-vibraciones, 2018.
- [56] Ángel Eduardo Rodríguez Suesca. *Evaluación de vibración de vibraciones en el diagnóstico de estructuras: Estudio de comportamiento de estructuras en deterioro*. PhD thesis, Departamento de Mecánica de Estructuras e Ingeniería Hidráulica, Universidad de Granada, Mayo 2021.
- [57] Ana Belén Roig Hernández. Medida y análisis de movimientos y vibraciones mediante sistemas de bajo coste basados en visión artificial, 2014.

- [58] Desineni Subbaram Naidu. *Optimal Control Systems*. CRC Press, 2003.
- [59] Donald E. Kirk. *Optimal Control Theory: An introduction*. Dover Publications Inc., 30 de Abril de 2004.
- [60] Katsuhiko Ogata. *Ingeniería de control moderna*. Pearson, 2010 quinta edición.
- [61] Fujita Kohei, Ayumi Ikeda, and Izru Takewali. Application of story-wise shear building identification method to actual ambient vibration. *Frontiers in Built Environment*, 2015.
- [62] M Lazarek, P Brzeski, and P Perlikowski. *Design and identification of parameters of tuned mass damper with inerter which enables changes of inertance*. Mechanism and Machine Theory, 2018.
- [63] Daniel J. Inman. *Vibration with Control*. John Wiley & Sons, 2006.
- [64] Dr. Reinaldo Welti. Vibraciones de estructuras complejas. *Ingeniería y Agrimensura de la Universidad Nacional de Rosario*, 2021.
- [65] Gobierno del Distrito Federal. Reglamento de construcciones para el distrito federal cgsservicios.df.gob.mx/prontuario/vigente/r38501.htm, 29 de enero de 2004.
- [66] Kevin Bonsor. The tuned mass damper: How science could earthquake-proof the skyscrapers of tomorrow science.howstuffworks.com/engineering/structural/smart-structure.htm, Octubre 2023.
- [67] Johan Antonio Pin Molina. Amortiguamiento de estructuras sites.ipleiria.pt/seismicknowledge/el-poder-del-pendulo-amortiguador-de-masa-sintonizado/, Agosto 2018.
- [68] Ing. Carlos Valdez. Disipadores de fluido viscoso como un sistema de protección sísmica para el diseño de edificios viconsconstructora.com/2021/06/28/disipadores-de-fluido-viscoso-como-un-sistema-de-proteccion-sismica-para-el-diseno-de-edificios/, 2021.
- [69] Johan Antonio Pin Molina. Amortiguamiento de estructuras sites.ipleiria.pt/seismicknowledge/amortiguadores-viscosos/, Agosto 2018.
- [70] Fayal Ikhouane and Jos Rodellar. *Systems with hysteresis: analysis, identification and control using the bouc-wen model*. Wiley-Interscience, 2007.
- [71] Montaña Moreno. *Redes neuronales artificiales aplicadas al análisis de datos*. PhD thesis, Universidad de Islas Baleares: Palma de Mallorca, España, 2002.
- [72] Damián Jorge Matich. Redes neuronales: Conceptos básicos y aplicaciones. Technical report, Universidad Tecnológica Nacional, Marzo del 2001.
- [73] Howard Demuth Mark Beale. *Neural Network Toolbox For Use with MATLAB*. The MathWorks, Octubre 2004.
- [74] Rivera Eduardo. Introducción a las redes neuronales artificiales. *Revista Científica No. 06 Universidad Don Bosco (UDB)*, Diciembre, 2005.

- [75] Pedro Isasi Viñuela and Inés M. Galván León. *Redes de Neuronas Artificiales Un enfoque práctico*. Pearson Prentice Hall, 2004.
- [76] Martin T. Hagan, Howard B. Demuth, and Mark Beale. *Neural Networ Design*. China Machine Press, 1996.