

PREDICCIÓN Y CATEGORIZACIÓN DE SERIES DE
TIEMPO USANDO UNA RED NEURONAL
MULTIFUNCIONAL

JONATHAN MAIRON RODRIGUEZ MARTINEZ
ANDERSON STIVEN QUINTERO ESPAÑOL

TRABAJO DE GRADO PRESENTADO COMO REQUISITO PARCIAL PARA OPTAR AL
TÍTULO DE INGENIERO ELECTRÓNICO.

DIRECTOR
IVAN DARIO LADINO VEGA
MAGISTER EN INGENIERÍA ELCTRÓNICA

FUNDACIÓN UNIVERSITARIA LOS LIBERTADORES
FACULTAD DE INGENIERÍA Y CIENCIAS BÁSICAS

BOGOTÁ D.C.
2022

NOTAS DE ACEPTACIÓN

Firma del presidente del jurado

Firma jurado

Bogota D.C Abril 2022

Agradecimientos

Quiero expresar principalmente mi gratitud a mi compañero de tesis Anderson Quintero por su apoyo dedicación y esfuerzo ya que su compromiso permitió realizar este trabajo de grado de manera satisfactoria.

A mis padres, amigos y esposa por confiar y creer en mis expectativas, por su apoyo, consejos y valores.

Agradezco de igual manera al Ingeniero Iván Ladino Vega quien ha sido mi instructor y guía para el desarrollo de este proyecto y que además fue un ejemplo a seguir a lo largo de mi trayectoria como estudiante, apoyándome con sus conocimientos y experiencia.

Finalmente agradezco a la Institución Universitaria los Libertadores y a todos los docentes por brindarme sus consejos y enriquecer mi conocimiento a lo largo de mi carrera.

Índice general

Agradecimientos	II
Lista de figuras	IV
1. Preliminares	2
1.1. Conceptos Clave	2
1.1.1. Serie de tiempo	2
1.1.2. Perceptrón	2
1.1.3. Red neuronal	3
1.1.4. Función de activación	4
1.1.5. Función de pérdida	5
1.1.6. Optimizador	5
1.1.7. Proceso de ajuste a una red neuronal	5
1.1.8. Convolución	6
1.1.9. Patrón Facade	6
1.2. Aclaraciones técnicas	7
1.3. Resumen	7
2. Redes neuronales para la categorización de ondas gravitacionales	8
2.1. Motivación	8
2.2. Introducción	8
2.3. Las series de tiempo generadas por las ondas gravitacionales	10
2.4. Redes neuronales convolucionales para el estudio de ondas gravitacionales	12
2.4.1. Especificaciones de la red implementada	13
2.5. Detección de parámetros	14
2.5.1. Generación de datos	14
2.5.2. Entrenamiento de la red	15
2.6. Resultados finales	15
3. Redes neuronales para identificación de irregularidades cardíaca	17
3.1. Motivación	17
3.2. Introducción	17
3.3. Las series de tiempo generadas por las pulsaciones cardíaca	18
3.4. Redes neuronales convolucionales para el estudio de electrocardiograma	19
3.4.1. Especificaciones de la red implementada	20
3.5. Identificación de irregularidades	21
3.5.1. Base de datos	21

3.5.2. Entrenamiento de la red	21
3.6. Resultados finales	22
4. Redes neuronales para la categorización de irregularidades cardíaca	23
4.1. Motivación	23
4.2. Introducción	23
4.3. Redes neuronales convolucionales para la clasificación de electrocardiograma	24
4.3.1. Especificaciones de la red implementada	24
4.4. Clasificación de irregularidades	25
4.4.1. Base de datos	25
4.4.2. Entrenamiento de la red	26
4.5. Resultados finales	26
5. Redes neuronales para identificación de movimientos sísmicos	28
5.1. Motivación	28
5.2. Introducción	28
5.3. Las series de tiempo generadas por los movimientos sísmicos	30
5.4. Redes neuronales convolucionales para la identificación de movimientos sísmicos	30
5.4.1. Especificaciones de la red implementada	31
5.5. Detección de sismos	32
5.5.1. Base de datos	32
5.5.2. Entrenamiento de la red	32
5.6. Resultados finales	33
6. Redes Neuronales estructuradas bajo el patrón de diseño Facade	34
6.1. Introducción	34
6.2. Desarrollo	35
6.2.1. Método de muestreo y entrenamiento	36
6.3. Resultados	37
7. Conclusiones	39
Bibliografía	40

Índice de figuras

1.1.	Separación binaria realizada por un perceptrón	3
1.2.	Separación no binaria realizada por una red neuronal	4
1.3.	Diagrama de Secuencia Patrón Facade. Imagen tomada de [9]	6
2.1.	Imagen de un agujero negro tomada por Event Horizon Telescope . . .	9
2.2.	Onda gravitacional generada por la colisión de dos agujeros negros. Tomada de [42]	10
2.3.	Onda gravitacional generada por la colisión de un objeto de 20 masas solares con otro de 35.	11
2.4.	Filtro de coseno y muestreo aplicado a una onda gravitacional	12
2.5.	Red Neuronal para la detección de parámetros	14
2.6.	Disposición de las masas	15
2.7.	Evolución del error absoluto medio en la detección de parámetros en ondas gravitacionales	16
3.1.	Fases de un electrocardiograma. Imagen tomada de [28]	19
3.2.	Red Neuronal para la detección de anomalías	21
3.3.	Evolución de la precisión del clasificador binario en función de las épocas	22
4.1.	Red Neuronal para la clasificación de afecciones cardíaca.	25
4.2.	Evolución de la precisión del clasificador multiclase en función de las épocas	27
5.1.	Red Neuronal para la predicción de sismos	31
5.2.	Precisión en la predicción sísmica	33
6.1.	Diagrama de distribución del redes neuronales con patrón Facade . . .	36
6.2.	Resultados en el entrenamiento del clasificador de dinámicas	37

Capítulo 1

Preliminares

1.1. Conceptos Clave

1.1.1. Serie de tiempo

El Análisis de Series de tiempo se ha convertido en una de las ramas más importantes y ampliamente utilizadas de la Estadística, sus campos de aplicación van desde la neurofisiología a la astrofísica y cubre áreas tan conocidas como el estudio de la previsión económica, los sistemas de control de datos biológicos, procesamiento de señales, telecomunicaciones y la física de vibraciones.

Una serie de tiempo es una secuencia de valores numéricos recopilados durante un periodo de tiempo, este conjunto de datos es en si la representación de un fenómeno físico, por ejemplo: Movimiento humano [3], ritmo cardíacos [4], el reconocimiento de fonemas [5], la predicción de terremotos, cantó de las ballenas y muchos más; Algunas definiciones formales para la clasificación de series de tiempo (TSC) pueden ser:

- **Serie de tiempo Univariante (UTS)** : Sólo una variable varía con el tiempo, Por ejemplo, datos recopilados de un sensor que mide la temperatura de una habitación cada segundo. Por tanto, cada segundo, solo tendrás un valor unidimensional, que es la temperatura y están representados por :

$X = [x_1, x_2, x_3, \dots, x_T]$ Conjunto ordenado de valores reales, la longitud de \mathbf{X} es igual al número de valores reales \mathbf{T} .

- **Series de tiempo Multivariable (MTS)** : Múltiples variables varían con el tiempo, por ejemplo, un acelerómetro Tri-axial hay tres aceleraciones, una para cada eje (x,y,z) y varían simultáneamente en el tiempo. Una serie de tiempo de M-dimensiones $X = [X^1, X^2, \dots, X^T]$ que contiene \mathbf{T} observaciones de $X^i \in \mathbb{R}^M$.
- **Base de Datos** : Una colección de datos $D = \{(X_1, Y_1), (X_2, Y_2), \dots, (X_N, Y_N)\}$ conformados por una conjunto de pares (X_i, Y_i) donde X_i podría ser una UTS o MST con Y_i como su vector de clase o etiqueta correspondiente.

1.1.2. Perceptrón

Un perceptrón [35] es un método de inteligencia artificial basado en la clasificación binaria supervisada. De tal manera, el algoritmo busca una función que, a partir de las variables de entrada, les asigne las etiquetas de salida esperadas.

Fue introducida por Frank Rosenblatt en 1957, quien buscaba replicar los principios básicos de la neurona biológica del ser humano.

El modelo utiliza una línea de hiperplano que clasifica dos entradas y las clasifica sobre la base de las 2 clases que aprende una máquina. La figura 1.1 representa un ejemplo de una clasificación hecha por un perceptrón.

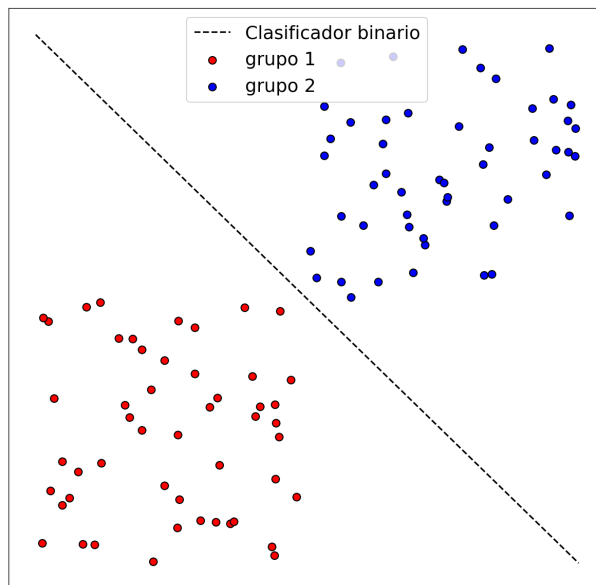


Figura 1.1: Separación binaria realizada por un perceptrón

1.1.3. Red neuronal

Una red neuronal [43] es la composición de capas en las que interactúan múltiples perceptrones, por lo que posee la capacidad de resolver problemas de clasificación y predicción más complejos que requieran de procesos de clasificación no binaria.

Cada conexión está asociada a un número llamado peso. La salida, h_i , de la neurona i , cada capa se define matemáticamente como:

$$h_i = \sigma \left(\sum_{j=1}^N V_{i,j} x_j + T_i^{hid} \right)$$

Donde σ es llamada la función de activación, N el número de entradas de la neurona, $V_{i,j}$ los pesos, x_j las entradas de las neuronas y T_i^{hid} el umbral de las neuronas ocultas. La innumerable cantidad de permutaciones para su construcción le permite ser capaz de resolver problemas no lineales tal y como se observa en la figura 1.2.

Gracias a la cantidad de cálculos y funciones que se deben ejecutar para su ajuste y ejecución, estos algoritmos son dependientes de los siguientes conceptos:

- Función de activación
- Función de pérdida
- Optimizador

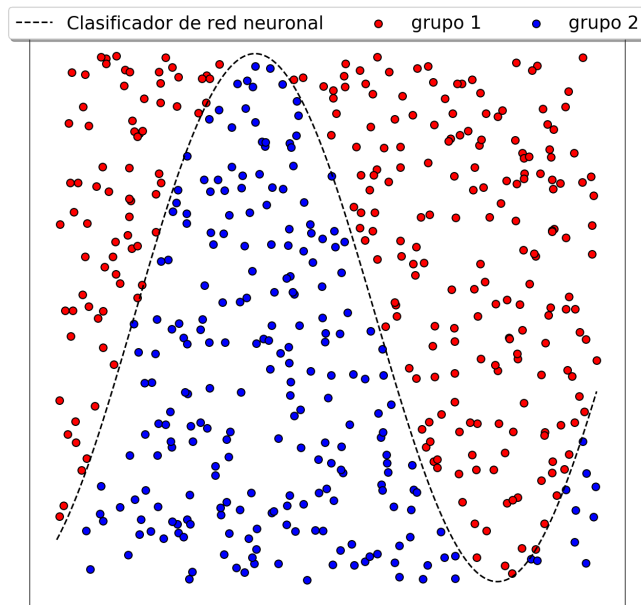


Figura 1.2: Separación no binaria realizada por una red neuronal

1.1.4. Función de activación

El propósito de la función de activación es [43], además de introducir la no linealidad en la red neuronal, limitar el valor de la neurona para que la red neuronal no se vea paralizada por neuronas divergentes.

Algunas de las funciones de activación más conocidas son:

Relu

La función de activación lineal rectificadora [13], o ReLU en su forma abreviada, es una función lineal a trozos que retornará la entrada si esta es positiva, de lo contrario, dará salida a cero. Se ha convertido en la función de activación por defecto para muchos tipos de redes neuronales porque un modelo que la utiliza es más fácil de entrenar y a menudo consigue un mejor rendimiento.

Softmax

La función Softmax es definida como:

$$\sigma(Z)_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}}$$

El término softmax [7] se utiliza porque esta función de activación representa una versión suave del modelo de activación .^{el} ganador se lo lleva todo.^{en} el que la unidad con la mayor entrada tiene una salida de +1 mientras que todas las demás unidades tienen una salida de 0.

1.1.5. Función de pérdida

El propósito de las funciones de pérdida es calcular la cantidad que un modelo debe tratar de minimizar durante el entrenamiento.

Algunas de las funciones de pérdida más utilizadas son:

Mean square error

El error cuadrático medio [26] se mide como la media de la diferencia al cuadrado entre las predicciones y las observaciones reales.

$$MSE = \frac{\sum_{i=1}^n (y_i - \bar{y}_i)^2}{n}$$

Sólo se refiere a la magnitud media del error. Sin embargo, debido al cuadrado, las predicciones que se alejan de los valores reales se ven muy penalizadas en comparación con las predicciones menos desviadas.

Entropía

Uno de los máximos exponentes en la teoría de la información es la conocida entropía [12], introducida por el llamado “Padre de la teoría de la información” Claude Elwood Shannon, en 1948, en el cual se establece la entropía como la incertidumbre media de la fuente de la información, matemáticamente se define como:

$$S = - \sum_{i=1}^W p_i \log_a(p_i)$$

Donde W es la cantidad de eventos posibles de la fuente de información Σ , p_i es la probabilidad de obtener un evento I_n .

1.1.6. Optimizador

Los optimizadores [17] son algoritmos utilizados para cambiar los atributos de la red neuronal, como los pesos y la tasa de aprendizaje, con el fin de reducir las pérdidas.

1.1.7. Proceso de ajuste a una red neuronal

Las redes neuronales al igual que gran parte de los algoritmos de inteligencia artificial requiere de un proceso de autocorrección que mejora su rendimiento sin una necesidad explícita de codificación.

Sin embargo, las redes neuronales son un tipo especial de modelos que pueden ser muy personalizables al contexto, por lo cual, todas las redes neuronales estandarizan un mismo proceso de corrección coloquialmente conocido como “entrenamiento”, el cual hace referencia al conjunto de procesos realizados para mejorar el ajuste de la red a los datos.

El proceso de corrección de resultados en una red neuronal es una compilación entre funciones de pérdida, para monitorear la disminución del error, un optimizador, que se encarga de encontrar el camino que minimiza el gradiente de error en la menor cantidad de iteraciones posibles y una función de activación, que se encarga de procesar la información obtenida por las neuronas.

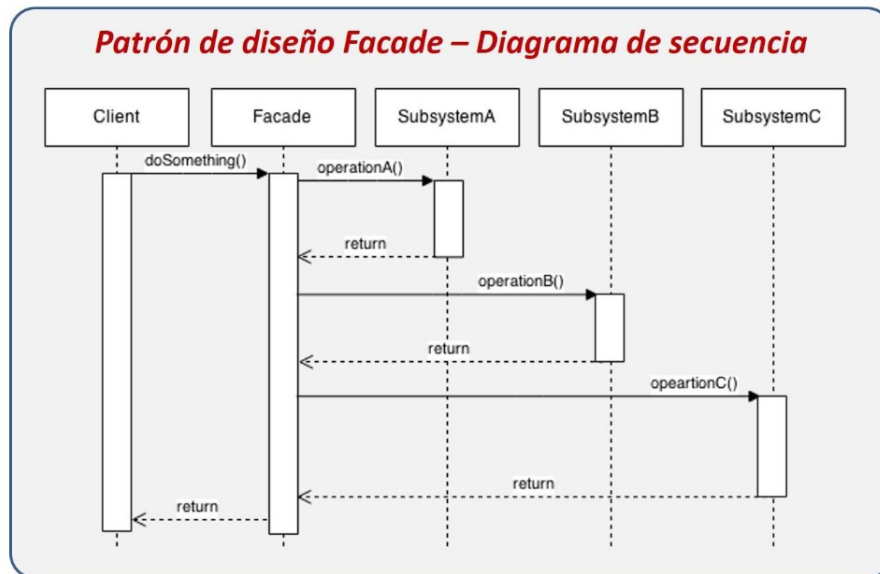


Figura 1.3: Diagrama de Secuencia Patrón Facade. Imagen tomada de [9]

1.1.8. Convolución

La convolución [40] es una operación que toma como entrada dos señales, una de ellas como la señal de entrada y la otra conocida como filtro sobre la imagen de entrada produciendo una señal de salida. La convolución es un concepto increíblemente importante en muchas áreas de las matemáticas y la ingeniería (incluida la visión por ordenador y el procesamiento de series de tiempo).

Considere entonces un vector de entrada f y un filtro g , donde f tiene longitud n y g tiene longitud m . La convolución $f * g$ se define matemáticamente como:

$$f * g = \sum_{j=1}^m g(j) f \left(i - j + \frac{m}{2} \right)$$

La convolución puede interpretarse como un promedio ponderado de la señal en la que los pesos son los valores del filtro.

1.1.9. Patrón Facade

El uso del patrón de diseño Facade hace uso de una interfaz de alto nivel que permite interactuar de forma mucho mas sencilla y ordenada con otros módulos de una aplicación, generalmente se implementa cuando existen fuertes dependencias entre clases y métodos que conforman un mismo sistema y se dice que es útil aplicarlo cuando una aplicación esta altamente acoplada y se requiere mejorar la comunicación e interacción reciproca con cada subsistema, el flujo ideal de secuencia se puede definir de la siguiente manera:

1.2. Aclaraciones técnicas

Todos los resultados presentados en este documento fueron obtenidos bajo codificación propia utilizando el lenguaje de programación Python en su versión 3.7. También se utilizaron algunas de sus librerías básicas como numpy, pandas, así como algunas más especializadas como sklearn, pycbc, keras y tensorflow. El entorno en el que se desarrollo se vario entre anaconda y Google Colab.

1.3. Resumen

Nuestro entorno constantemente nos rodea con información, desde pequeños detalles como síntomas en el interior de nuestro cuerpo que identifican un mal funcionamiento en el organismo, hasta un leve cambio de temperatura que predice una futura lluvia. La naturaleza constantemente genera datos que no siempre podemos detectar debido a las limitaciones de nuestros sentidos y la capacidad específica de razonamiento que no esta diseñada para almacenamiento y procesamiento de grandes flujos de información. En este trabajo exploraremos 4 fenómenos de diferentes áreas, biológicos, físicos y naturales las cuales poseen varias características en común que los relacionan, la mas destacable, es que todas pueden ser representadas cuantitativamente usando series temporales. Sin embargo, a lo largo de este documento veremos como estas 4 dinámicas poseen características tan similares que les permiten ser estudiadas, caracterizadas e incluso, se puede extraer información valiosa de ellas implementando un mismo algoritmo.

El algoritmo utilizado es una estructura de red neuronal convolucional de creación propia especializada en el estudio de series de tiempo para el análisis, detección, clasificación y asignación de tareas basadas en series de tiempo que no tiene la limitación estricta de entrenar usando gigantescos flujos de información, la cual es una de las restricciones principales en los modelos de Deep learning. Esta red fue basada en los algoritmos utilizados para el reconocimiento de fenómenos astrofísicos inyectados en series de tiempo altamente ruidosas.

Además, a lo largo del trabajo se evaluará las diferentes limitaciones de la red, así como Identificar en qué condiciones es adecuado utilizarla y en que posibles fenómenos también podría utilizarse. Finalmente los algoritmos se acomodan aplicando el patrón de diseño Facade con el fin de estructurar el conjunto de redes neuronales como una aplicación escalable que permita un crecimiento horizontal, es decir, que permita agregar mas redes neuronales que cumplan con los criterios de aceptación (series de tiempo).

El resultado es una clasificación y solución de situaciones que requieren de análisis de series temporales que no pueden ser solucionadas a partir de los sentidos humanos, siendo este un pequeño vistazo a inteligencias artificiales interactivas a las necesidades humanas.

Capítulo 2

Redes neuronales para la categorización de ondas gravitacionales

*El alma nunca piensa sin una imagen
mental*
Aristóteles

2.1. Motivación

La astrofísica se ha encargado de estudiar eventos astronómicos que ocurren a millones de kilómetros de la tierra e identificar sus dinámicas. Uno de los fenómenos más intrigantes es la colisión de objetos cuya densidad supera a la del sol, para esto se estudia la energía residual del impacto la cual se propaga como ondas a través de todo el espacio y que comúnmente son opacadas por ruido espacial.

En la actualidad se han realizado grandes avances en el campo de la extracción de señales inyectadas en ruido, por lo que al extraer una aproximación de la señal se logra obtener grandes fuentes de información de las masas que las generaron.

De tal manera, el objetivo de esta sección es analizar un tipo especial de series de tiempo, las ondas gravitacionales y a partir de ellas detectar los objetos que generaron estas perturbaciones utilizando una red neuronal convolucional basada en las redes convencionalmente usadas en la literatura y simplificándolas para tener mejores rendimientos sin sacrificar la precisión de la detección.

2.2. Introducción

El campo experimental de la física se ha enfocado en describir y aprender sobre fenómenos cuyo comportamiento aun es desconocido por la humanidad, este particular interés ha contribuido en innumerables avances para el desarrollo tecnológico [20], a su vez que incrementa el entendimiento que tenemos sobre el habitad que nos rodea.

En la actualidad se han desarrollado herramientas altamente avanzadas que permiten explorar ambientes a los que difícilmente un humano podría acceder, un ejemplo de esto son las contribuciones en el campo de la astrofísica, cuyo eje de estudio se centra en la observación de masas inconmensurables y las interacciones entre estas.

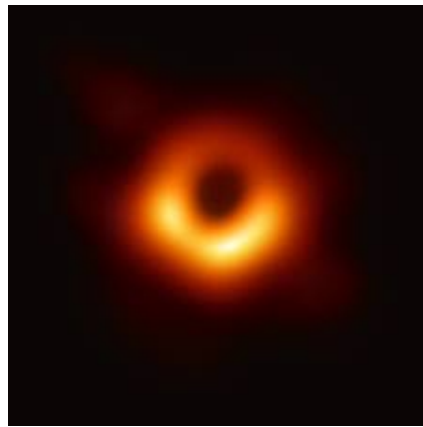


Figura 2.1: Imagen de un agujero negro tomada por Event Horizon Telescope

Una masa inconmensurable [34] se suele definir como un objeto de gran tamaño que iguala o incluso supera el tamaño del sol. Para categorizar estos cuerpos se usa el sol como unidad base de medida, por lo que objetos de n masas solares hacen referencia a cuerpos cuyas masas son n veces más grandes que la del sol.

Existen diferentes masas inconmensurables, sin embargo, algunas de las más populares son los agujeros negros [21], objetos lo suficientemente densos como para generar una atracción gravitacional tan fuerte que ni siquiera la luz escapa de ella. Estos objetos son especialmente importantes ya que fueron descritos hipotéticamente por Einstein en su teoría de la relatividad general hasta que fue corroborada años después.

Dada la naturaleza de los cuerpos, las masas de tamaño inconmensurable se ubican a distancias muy lejanas de la tierra, por lo que garantizar la existencia de objetos como agujeros negros es un desafío altamente complejo incluso utilizando equipo especializado que discrimine objetos como estrellas o ruido espacial de los cuerpos de interés.

Sin embargo, pese a las dificultades que presenta el asegurar la existencia de masas inconmensurables, en el año 2015 [25] se logró detectar una, a partir de un concepto conocido como onda gravitacional.

Las ondas gravitacionales son propagaciones de energía que viajan a través del espacio, las cuales son producidas por la colisión de dos o más objetos lo suficientemente densos como para deformar fuertemente el espacio tiempo. Estos fenómenos ondulatorios producen un tipo especial de series, ondas que pueden ser detectadas a partir de instrumentos de medición altamente precisas en conjunto de herramientas de detección autónoma [6].

Por lo general gran parte de estas perturbaciones son producidas por agujeros negros colisionando y generando una nueva masa con una compacidad superior de las cuales se originó. Sin embargo, teóricamente las ondas gravitacionales también podrían ser generadas por otras interacciones, de la mas destacables se encuentran remanentes del big bang, la famosa teoría que afirma que todo lo conocido en el universo parte de una explosión. Hipotéticamente se estima que residuos energéticos de dicha explosión aún se encuentran en el espacio tiempo incluso hoy en día.

De tal manera, el interés principal por detectar ondas gravitacionales es identificar la fuente de origen de esa energía y estudiar a profundidad las propiedades de este objeto. Por lo tanto, con lo anterior establecido, en esta sección se realiza un estudio de algunas series de tiempo obtenidas por las ondas gravitacionales con el fin de establecer carac-

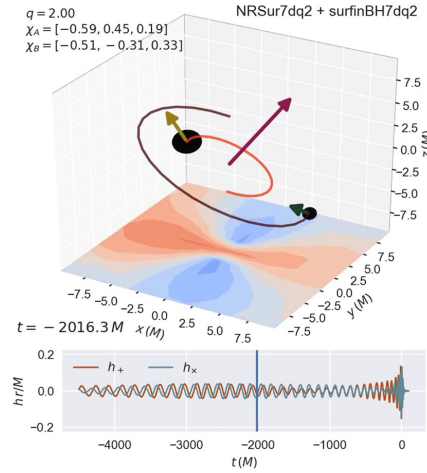


Figura 2.2: Onda gravitacional generada por la colisión de dos agujeros negros. Tomada de [42]

terísticas de las masas que las generaron e identificar que tipo de masas produjeron el fenómeno.

Para esto, se mostrará explícitamente las ondas gravitacionales vistas a partir de una serie de tiempo, como se procesaron, el origen de la base de datos que se estudio y el resultado de aplicar la red neuronal en el reconocimiento de parámetros.

2.3. Las series de tiempo generadas por las ondas gravitacionales

El comportamiento de las ondas gravitacionales posee una evolución particular debido al entorno libre de fricción en el que se genera la dispersión de la energía y las particularidades de las masas que interactúan.

A diferencia de las ondas usadas para describir la posición de un resorte o la caída libre de un objeto, la colisión de masas actúa de forma pseudo periódica en el que se incrementa la amplitud de sus ondas en cada periodo, esto se debe a que al pasar el tiempo la cercanía entre los objetos es mayor, lo que se traduce en una atracción mas fuerte entre los objetos.

La figura 2.3 muestra la evolución de una onda gravitacional generada a partir de dos objetos de 20 y 35 masas solares y se puede observar como a medida que el tiempo pasa, el incremento de la amplitud es progresivo hasta llegar a un punto máximo, el cual se muestra en el primer acercamiento marcado resaltado con rojo de la figura 2.3. Este es el momento en donde las masas colisionan y generan un nuevo cuerpo.

Las fases posteriores a este encuentro son residuos de la energía posterior a la colisión, por lo que progresivamente decrentan la propagación de energía hasta llegar a un punto de estabilidad, este comportamiento es evidenciado en el segundo acercamiento marcado con azul de la figura 2.3.

Como se ilustro anteriormente, la amplitud máxima, el momento de estabilización y la frecuencia de las ondas, generan información importante sobre el fenómeno, características que permitirían las masas que las generaron, ya que, de ser una colisión entre objeto con una menor diferencia entre masas, la frecuencia sería menor, el punto máximo de colisión tendría una menor amplitud y el momento de estabilización se podría

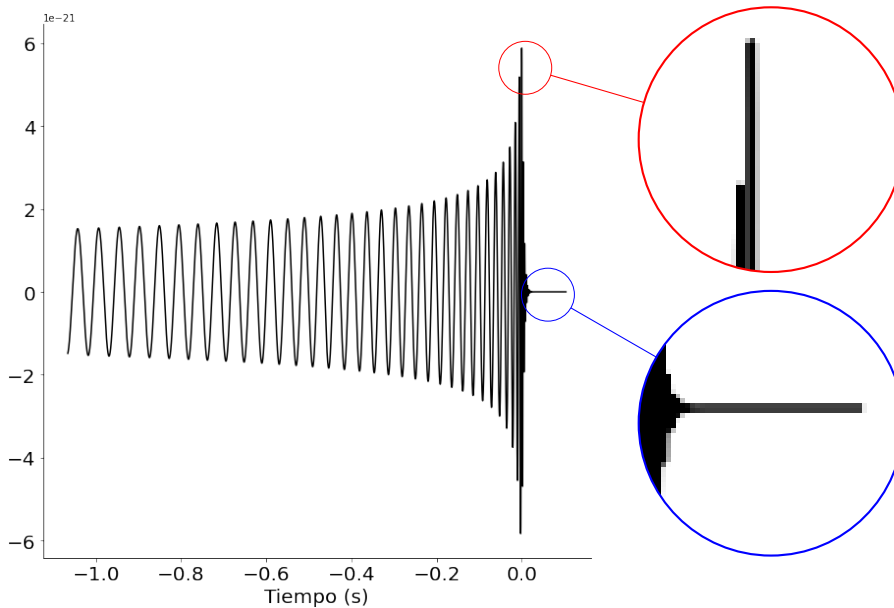


Figura 2.3: Onda gravitacional generada por la colisión de un objeto de 20 masas solares con otro de 35.

ver reducido. Análogamente, los objetos con una mayor diferencia de masas generar movimientos con una menor frecuencia, mas tiempo de estabilidad según sea el caso y un punto de perturbación máximo mucho mayor.

Por lo tanto, se estima que al destacar estas características se podría generar un estudio mas preciso que permita reconocer con una alta precisión la magnitud de las masas sin recurrir a herramientas altamente complejas que requieran de un coste computacional muy elevado.

Si bien, visualmente se pueden identificar el punto máximo y la fase de estabilidad, para poder generar un algoritmo que detecte estas particularidades es necesario resaltarlos individualmente, de modo que la información trivial en la serie de tiempo sea opacada y fácilmente procesada.

Dado a que cada colisión depende de una permutación de masas, las dinámicas máximas de propagación de las ondas llegan a variar en la posición y amplitud, por lo que detectar estas características es un proceso que requiere analizar la serie de tiempo individualmente, lo que en términos computacionales es ineficiente.

Para dar solución al reconocimiento, el resaltado de características se realiza mediante el proceso de convolución, el cual, a partir de un filtro o función, se resaltan características de interés para una serie de tiempo, en específico, para las ondas gravitacionales las convoluciones se presentan como una herramienta muy efectiva para el análisis de series de tiempo [45].

Gracias a los filtros, existe una herramienta que destaca características en las series de tiempo, sin embargo, para el caso de las ondas gravitacionales, la mayor parte de las oscilaciones son menos relevantes a las que se genera en el pico máximo y en el reposo final del comportamiento, por lo cual, generar un muestreo posterior al filtro no solo disminuye la cantidad de datos con las que se procesa, sino que discretiza las regiones

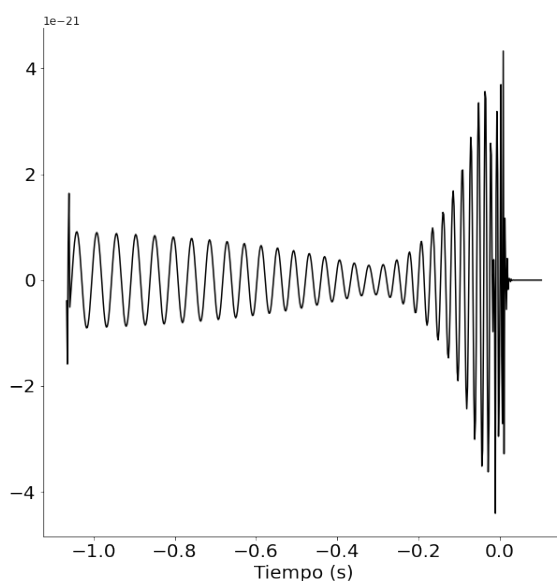


Figura 2.4: Filtro de coseno y muestreo aplicado a una onda gravitacional

de interés en el fenómeno para analizar.

La figura 2.4 muestra una simplificación resaltada de la señal 2.3 a la que se le aplicó un filtro del coseno en base a una señal de 64 términos. Posteriormente la señal filtrada fue muestreada, siendo la serie de tiempo original 2.3 un conjunto de 4791 elementos, que al simplificar se obtuvo 2.4 de 599 términos.

Como se puede observar en la figura 2.4, una convolución unidimensional puede generar resalte de características a las ondas gravitacionales manteniendo aspectos como el punto máximo de energía y la fase de reposo disminuyendo la pérdida de información relevante a partir del muestreo. Lo que facilita el proceso de análisis general sin la pérdida perceptible de información relevante.

Sin embargo, no es posible garantizar que el filtro de coseno o algún otro filtro de resalte convencional permitan detectar de manera eficiente los parámetros de las masas. Dada la innumerable cantidad que puede implementarse, es necesaria una herramienta que ajuste y adapte un filtro a las características específicas del estudio y lo enfoque en detectar ondas gravitacionales.

2.4. Redes neuronales convolucionales para el estudio de ondas gravitacionales

Actualmente no existe un método explícito o una expresión algebraica general capaz de identificar el mejor filtro para cada caso puntual, por lo que el reconocimiento de filtros se suele basar en un sistema de autocorrección alimentado a partir de un proceso de prueba y error. Esto implica que identificar un algoritmo de reconocimiento que se base en la experimentación continua y que continuamente mejore de acuerdo con la experiencia adquirida a partir de metodologías fallidas.

La descripción anterior se acopla adecuadamente a un método de inteligencia artificial, ya que, para poder realizar un análisis adecuado, estos algoritmos especializados analizan, comparan y desarrollan cierto “aprendizaje” que les permite crear “experiencias”, simulando así el reconocimiento de patrones de forma análoga a como lo haría un ser

humano.

En específico, los algoritmos de acoplamiento artificial más reconocidos, tanto por la personalización que ofrecen como en la flexibilidad de sus predicciones y clasificaciones son las redes neuronales [1].

Ya que sus características permiten resolver problemas muy sencillos o complejos si se tiene la cantidad de datos suficientes y si se logra determinar una estructura adecuada basada en el desarrollo del fenómeno. Las redes neuronales han demostrado tener gran potencial para resolver problemas de todo tipo de tópicos como la detección de cáncer [41], la conducción autónoma [27] o incluso el estudio de ondas [31].

A demás, Las redes neuronales se acoplan fácilmente a las convoluciones las cuales generan una nueva categoría de red neuronal, las redes neuronales convolucionales (CNN).

Las (CNN) son un tipo especial de redes neuronales que además de poseer una estructura basada en perceptrones compuestos, se le es agregada un conjunto de parámetros con el objetivo de realizar convoluciones y que al igual de las neuronas, serán ajustados iterativamente a partir del entrenamiento.

2.4.1. Especificaciones de la red implementada

Específicamente, para el procesamiento y detección de ondas gravitacionales la literatura suele utilizar redes neuronales altamente profundas con varias capas de convolución y un aproximado de 4 capas ocultas para realizar los estudios tal y como se muestra en [32, 39].

Esto se debe a que las señales con las que se suelen trabajar previamente están ocultas en altas intensidades de ruido, tal como se muestra en [39], sin embargo, debido a los avances que han tenido la extracción de señales de fuentes ruidosas [38, 32] es posible generar un estudio de series de tiempo tomando como entrada la onda gravitacional pura sin interacciones de ningún tipo, reduciendo así la necesidad de capas altamente profundas para la extracción de características alteradas.

Por otro lado, para procesar los datos de las ondas usualmente se trabajan en series de tiempo de 8192 datos [44, 32], sin embargo, debido al coste computacional que requiere este procesamiento, en contraposición se utilizará un muestreo de estas series de 2190, el cual mantendrá las características de las series originales, pero utilizando una recolección en segundos mucho menor.

De tal manera, las entradas en esta red son vectores de 2192 elementos, que serán pasados a partir de 8 filtros de longitud 64 que serán aplicados corriendo dos pixeles en lugar de uno, en otras palabras, tendrá un stride de 2 y una función de activación relu. La razón para aplicar una intensidad de filtros de tal magnitud es porque las partes más relevantes de las ondas son el punto máximo y la fase de reposo, las cuales son de interés para el reconocimiento de parámetros.

Posteriormente el resultado de la convolución pasara a dos capas densas de 64 y 32 neuronas respectivamente para finalmente obtener un vector de dos elementos activado a partir de la función de activación relu. Estos dos parámetros finales corresponderán a las masas que colisionaron para generar la onda gravitacional, tal como lo ilustra la figura 2.5.

La característica principal de esta red es que posee muchos menos pesos de entrenamiento, lo que permite tener una menor cantidad de valores en la serie de tiempo, requiere menos tiempo de entrenamiento, y esta mas enfocada en la caracterización de ondas gravitacionales.

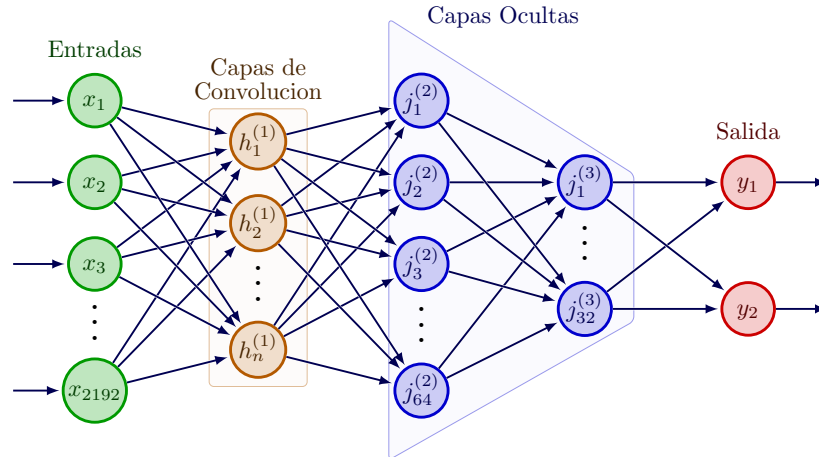


Figura 2.5: Red Neuronal para la detección de parámetros

De tal manera, el principal aporte de esta red es generar una detección de parámetros a partir de una red neuronal relativamente sencilla que al ser optimizada permita realizar detecciones de bajo costo computacional y funcione en dispositivos de bajo almacenamiento y que sea fácilmente replicable para cualquier persona sin la necesidad de ordenadores de alto rendimiento.

Finalmente, en la siguiente sección se explica el origen de los datos utilizados, la metodología para entrenar la red neuronal y los resultados finales de la red.

2.5. Detección de parámetros

2.5.1. Generación de datos

En la actualidad, es altamente complejo obtener los datos reales sobre ondas gravitacionales, ya que no es usual captar este tipo de señales y mucho menos obtenerlas puras, por lo que es usual generar datos sintéticos para los estudios relacionados con estos fenómenos.

Para este objetivo en este trabajo se implemento aproximaciones a estos comportamientos a partir de detecciones previamente conocidas, de esto se encarga la librería PyCBC [8], el cual es un paquete de software libre utilizado para explorar fuentes astrofísicas de ondas gravitacionales.

Sin embargo, la librería genera las ondas gravitacionales a partir de dos parámetros correspondientes a las masas de los objetos, por lo que fue necesario generar una distribución de estas, para ello se generaron 703 parejas de masas que oscilaban entre las 10 masas solares y las 100 masas solares y se dividieron en 492 para el entrenamiento, 106 para probar el algoritmo y 105 para validarlo. Se distribuyeron como se muestra en la figura 2.6.

Cada punto en la imagen 2.6 representa dos masas colisionando y generando una onda gravitacional, de tal manera, estas masas serán utilizadas en la librería PyCBC con el fin de generar series de tiempo necesarias para ajustar el algoritmo. Las series de tiempo fueron utilizadas como entradas del algoritmo y las dos masas asociadas son las etiquetas que se tomarán como respuesta de la red.

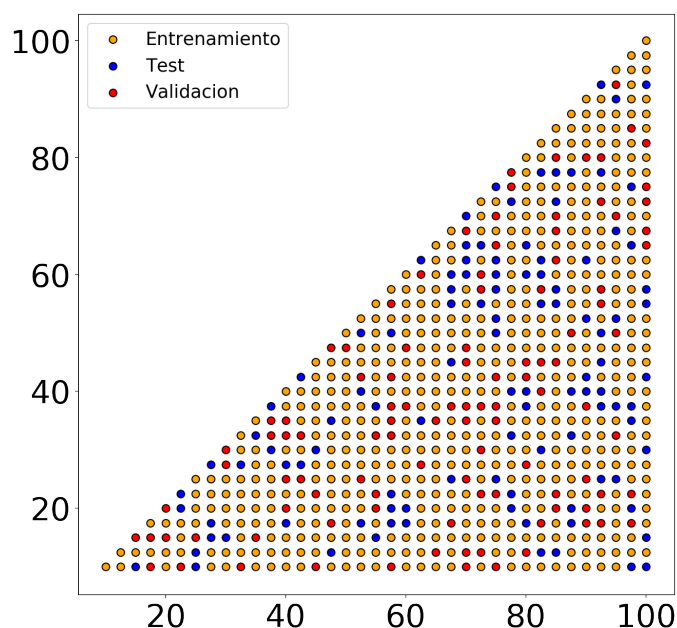


Figura 2.6: Disposición de las masas

2.5.2. Entrenamiento de la red

Para ajustar la red a los datos se utilizó como función de pérdida el error medio cuadrático, sin embargo, también se monitoreó utilizando el error absoluto medio para facilitar la comprensión del error, como optimizador se utilizó el método Adam.

Para el entrenamiento se utilizaron 1000 épocas implementando las ondas destinadas de prueba para evaluar el error a los datos desconocidos por el algoritmo.

Finalmente, los resultados del entrenamiento de la red y el rendimiento de esta ante los datos desconocidos durante su ajuste serán mostrados en la siguiente sección.

2.6. Resultados finales

En esta sección se evalúan los resultados de la red en la fase de entrenamiento y el error de la red ante los datos de validación. En su fase de ajuste, la red posee un decrecimiento progresivo del error absoluto medio, tanto en las predicciones utilizadas por la red al entrenar como las pruebas ejecutadas para minimizar su error a los datos desconocidos tal y como se muestra en la figura 2.7.

Existen algunas variaciones en el decrecimiento progresivo del error que indican la exploración de la red a otra distribución de neuronas, sin embargo, pese a la exploración generada el error regresa a puntos de estabilidad constante y monótono tanto en el entrenamiento (línea azul) como en la prueba (línea roja).

Los resultados indican que el error medio de la predicción está entre 3 y 5 masas solares, lo cual es una aproximación lo suficientemente precisa para una detección de parámetros preliminar.

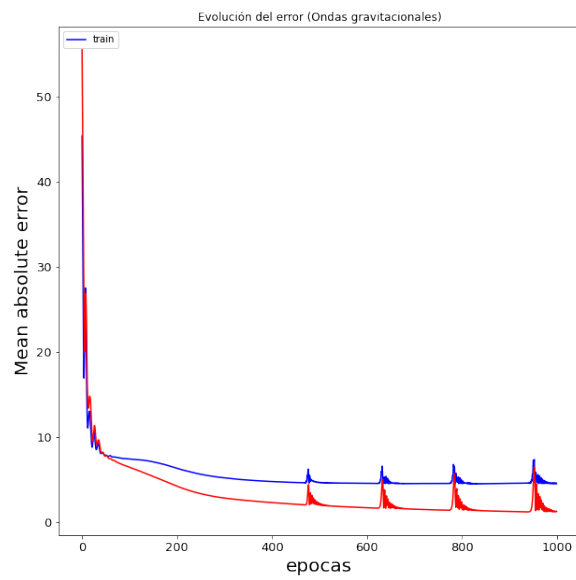


Figura 2.7: Evolución del error absoluto medio en la detección de parámetros en ondas gravitacionales

Capítulo 3

Redes neuronales para identificación de irregularidades cardíaca

*El alma nunca piensa sin una imagen
mental*
Aristóteles

3.1. Motivación

La ciencia de la salud ha generado enormes beneficios en la esperanza de vida mundial, el trabajo de los profesionales permite establecer soluciones eficientes y premeditadas con el fin de enfrentar peligrosas enfermedades sin sufrir consecuencias mucho mayores.

Sin embargo, en ocasiones las señales que le permiten al profesional identificar un diagnóstico son poco perceptibles que desencadenan errores en los diagnósticos, esta limitación puede verse concretamente en las afecciones cardíaca, las cuales demandan que el médico detecte cambios irregulares en los latidos del corazón que pueden tardar menos de un segundo.

De tal manera, el objetivo de esta sección es generar un clasificador binario que determine con alta eficiencia cuando un corazón no presenta un comportamiento natural, para ello se analizaran un conjunto de series de tiempo asociadas al corazón, los electrocardiogramas, que se tomaran como datos de entrenamiento en la red neuronal anteriormente implementada para generar una dupla de probabilidades que indiquen con alta precisión si un corazón presenta una irregularidad o no.

3.2. Introducción

Además de implementar la tecnología para la exploración de entornos y el entendimiento de la naturaleza, la inteligencia artificial ha permitido generar valiosas herramientas para incrementar la calidad de vida de las personas, desarrollos como la conducción autónoma, la visión por computadora son ejemplos de estos avances.

Sin embargo, entre los aportes mas valiosos en cuanto al bienestar general se trata es la detección temprana de enfermedades, ya que los malestares que aquejan al ser hu-

mano son difícilmente identificables incluso para los especialistas más experimentados y reconocer de manera efectiva estas implicaciones es crucial para poder aplicar un tratamiento efectivo.

Existen antecedentes mundialmente reconocidos de inteligencias artificiales implementadas para el apoyo de la medicina [16], como en la detección temprana de tumores cancerígenos, diagnóstico de Alzheimer, segmentación de imágenes cerebrales y en específico, la detección de irregularidades mediante el análisis de señales.

Es sabido que los métodos convencionales para la detección de enfermedades como deficiencias en la arteria coronaria es costoso [16], además de presentar muchos inconvenientes para el paciente, así como presentar grandes tiempos de análisis, tiempo del cual no se suele contar.

Con base a esto, el estudio de las acústicas cardíaca es de gran interés para el campo de la medicina, ya que permitiría segmentar pacientes irregulares y generar un mayor estudio de sus afecciones mientras que se disminuyen los costos de su posterior diagnóstico

No obstante, realizar este estudio no es un proceso sencillo, ya que el bombeo y la frecuencia cardíaca de cada persona varía de acuerdo con los hábitos de vida que se practiquen individualmente, así como las características genéticas particulares del individuo en cuestión, por lo que tiende a ser un proceso que requiere de métodos de identificación de patrones especializados totalmente adaptados a este fenómeno.

A partir de lo anterior, esta sección se centrará en explicar las series de tiempo generadas para el estudio del corazón, los electrocardiogramas y como estas series de tiempo pueden ser procesadas para generar información útil para el tratamiento de los pacientes, en específico identificar las personas que tengan una irregularidad en sus latidos.

3.3. Las series de tiempo generadas por las pulsaciones cardíaca

Pese a los limitados movimientos del corazón, las señales emitidas por el este generan una gran cantidad de información que permiten monitorear adecuadamente las condiciones del órgano y su correcto funcionamiento. Las instituciones médicas implementan un equipo especializado conocido como electrocardiógrafo cuya función es a partir de una prueba no invasiva registrar gráficamente la actividad eléctrica del corazón. Las señales representadas por el equipo se conocen como electrocardiograma.

Debido a que la única interacción fisiológica del corazón con el cuerpo humano es el bombeo de sangre a partir de su contracción y expansión, las representaciones del electrocardiograma sobre el papel suelen verse como ondas en las que se identifican el ritmo y la regularidad de los latidos, el tamaño y posición de las aurículas y ventrículos, cualquier daño al corazón y los efectos que sobre él pueden tener ciertos fármacos o dispositivos implantados en el corazón (como marcapasos).

Como se ilustra en la imagen 3.1, se observa como las fases de las pulsaciones representan partes esenciales del movimiento cardiaco y como este genera diferentes interacciones en cada parte del corazón

- La onda P es un registro de la actividad eléctrica a través de las cavidades superiores del corazón (aurículas) [15].
- El complejo QRS es un registro del movimiento de los impulsos eléctricos a través de las cavidades inferiores del corazón (ventrículos) [15].
- El segmento ST muestra cuando se contrae el ventrículo, pero no hay electricidad fluyendo a través de él. El segmento ST generalmente aparece como una línea

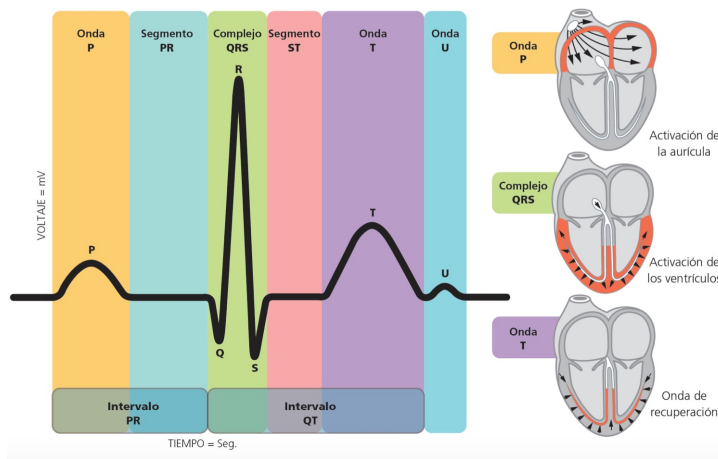


Figura 3.1: Fases de un electrocardiograma. Imagen tomada de [28]

recta horizontal entre el complejo QRS y la onda T [15].

- La onda T muestra el momento en que las cavidades inferiores del corazón se restablecen eléctricamente y se preparan para la siguiente contracción muscular [15].

Estas fases pueden tener pequeñas variaciones dependiendo de la edad y de la composición física de las personas, sin embargo, para un corazón saludable se tiene previamente estructurado una evolución general ideal de la composición de la onda, la cual determina de ser totalmente representada en un electrocardiograma se determina si un individuo es saludable o no con gran exactitud.

Sin embargo, el principal problema para determinar la naturaleza de un corazón sano radica en que a veces las irregularidades en las ondas son tan pequeñas que el ojo humano no logra detectarlas con facilidad.

Por lo cual, las metodologías de autocorrección de las redes neuronales se muestran como una herramienta altamente eficiente para la detección temprana de enfermedades cardíacas, en evidencia a esto, la literatura establece un precedente de efectividad de esta metodología en monitorear y acoplar redes neuronales al comportamiento del corazón [18, 36, 30].

3.4. Redes neuronales convolucionales para el estudio de electrocardiograma

Se estima que para un cuerpo en reposo y totalmente saludable un electrocardiograma será una señal totalmente periódica, por lo que para un algoritmo computacional no sería complicado detectar un comportamiento irregular si se establece una base guía adecuada de comparación, sin embargo, las pequeñas variaciones de edad y tamaño podrían generar un margen de error importante que es necesario acoplar en el estudio. De manera análoga al caso de las ondas gravitacionales, los puntos de relevancia de la onda se enfocan en partes específicas del comportamiento, en este caso los segmentos PR, el complejo QRS y la onda T por lo que específicamente los intervalos de pausa

entre señales son menos relevantes en la caracterización de irregularidades, de manera que al simplificar la señal mediante procesos de convolución destacara las características requeridas haciendo más fácil la detección.

A grandes rasgos, las características de ondulatorias de las pulsaciones del corazón poseen gran similitud con las ondas gravitacionales, ya que la mayor fuente de información se centra en puntos específicos de las ondas, que al pasar por el proceso de convolución son resaltadas, mientras que los segmentos en los que no hay actividad son opacados por la amplitud resaltada de los eventos más representativos, de tal manera, en la siguiente subsección se tratara como se acoplo la red utilizada para determinar los parámetros de las ondas gravitacionales para establecer un clasificador binario que determina si un corazón posee irregularidades o no.

3.4.1. Especificaciones de la red implementada

Esta red es un clasificador binario, en otras palabras, se busca detectar si el comportamiento del corazón analizado es natural o si posee algún comportamiento irregular, de tal manera, dada la naturaleza del fenómeno, especificar de forma certera el comportamiento de los latidos cardiacos es dependiente a las características fisiológicas del paciente, debido a esta cierta incertidumbre, es necesario asignar un cierto grado de incertidumbre a las detecciones según la evidencia obtenida.

Para ello, la literatura recomienda utilizar como métrica un concepto enfocado en la irregularidad de los datos, la entropía [2, 3], cuyos resultados en la detección han generado enormes resultados en procesos donde no basta solo con el conocimiento de la red, sino que pueda ser validado y corroborado por el criterio humano.

De manera que los resultados obtenidos por la red serán un conjunto de dos probabilidades que indican la medida de certeza que se tiene cada categoría, la probabilidad de que sea un corazón saludable o por el contrario tenga una afección.

A diferencia de las ondas gravitacionales, el periodo de las ondas solo es afectado cuando se presenta el caso de irregularidades, lo que provee de un gran apoyo a la red y a las irregularidades finales en la convolución.

Las entradas de esta red son series de tiempo dispuestas en forma vectorial de 108 valores, que serán procesados a partir de la red que se utilizo para detectar los parámetros de las ondas gravitacionales, en otras palabras se utilizara una red con las siguientes especificaciones: 8 filtros de longitud 64 que serán aplicados cada dos pixeles, activados mediante la función relu, para posteriormente pasar a dos capas densas de 64 y 32 neuronas respectivamente, la estructura de esta red puede observarse en la figura 3.2.

La función de perdida utilizada es la entropía y la métrica de evaluación es la precisión o accuracy, que se calcula dividiendo la cantidad de casos favorable o clasificados correctamente, sobre la cantidad de casos totales.

Se espera que de manera análoga a como ocurrió en las ondas gravitacionales, la detección de las ondas debido a las convoluciones y la naturaleza de las ondas generadas en la serie de tiempo sean comportamientos altamente predecibles.

Por lo tanto, el segundo aporte de este estudio además de la poca cantidad de parámetros utilizado para la detección y categorización es la replicabilidad a fenómenos de características ondulatorias.

Finalmente, en la siguiente sección se explica el origen de los datos utilizados, la metodología para entrenar la red neuronal y los resultados finales de la red.

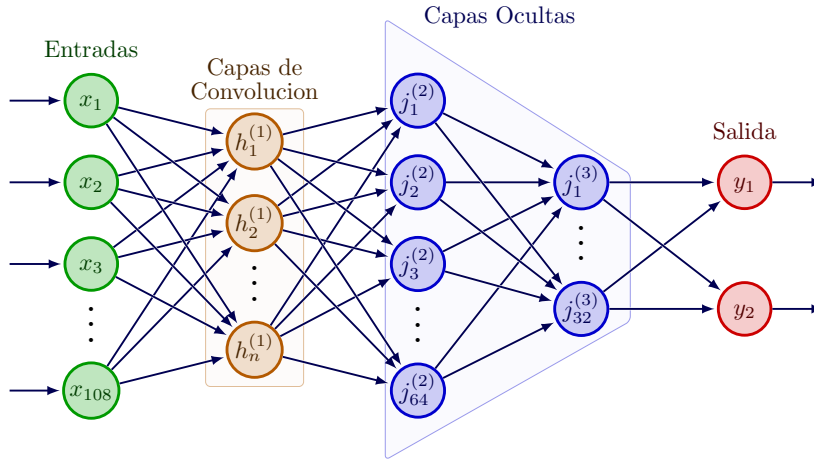


Figura 3.2: Red Neuronal para la detección de anomalías

3.5. Identificación de irregularidades

3.5.1. Base de datos

La base de datos obtenida de [19] y esta dispuesta en dos secciones, la primera contiene 4046 registros de electrocardiogramas tomados de corazones con latidos normales, mientras que la segunda contiene 10506 series de latidos anormales. Cada registro es una serie de tiempo de 188 elementos que fue muestreada a una frecuencia de 125 HZ. Los datos de los cuales se basa el repositorio [19] están basados en [23, 11], cuyos registros eran tomados de sujetos de 17 a 87 años, con una media de 57.2; 209 hombres, con una media de 55.5 años, y 81 mujeres, con una media de 61.6 años. Sin embargo, la descripción de los datos de [19] no estipula esta información explícitamente.

3.5.2. Entrenamiento de la red

Para el entrenamiento se utilizaron los 14552 registros disponibles etiquetados a partir de vectores categóricos de dos componentes, para el caso de que un registro fuera de un corazón sin alteraciones el vector de etiqueta sería de la forma $(1, 0)$ indicando que la probabilidad de que el individuo no tenga alteraciones es de 1, análogamente si la serie de tiempo representa algún comportamiento con afecciones la etiqueta asociada será de la forma $(0, 1)$

Los datos se distribuyeron en tres conjuntos, 70% del total, lo que equivale a 10186 series de tiempo entrenaran la red, mientras que el resto será distribuido de forma equivalente entre el conjunto de datos de testeo y validación, esto indica que hay 2182 series para cada conjunto.

Como optimizador se utilizó el método Adam, como métrica de evaluación se implementó la precisión o accuracy y como función de pérdida se utilizó la entropía binaria. En el análisis preliminar del modelo se dejaron interactuar diferentes épocas hasta que se determinó que la convergencia del modelo converge adecuadamente después de las 80 épocas, por lo que se le asignaron 100 épocas de entrenamiento

Finalmente, los resultados del entrenamiento de la red y el rendimiento de esta ante los datos desconocidos durante su ajuste serán mostrados en la siguiente sección.

3.6. Resultados finales

Una vez entrenado el modelo binario, se puede apreciar la evolución de la precisión en la figura 3.3, como se puede apreciar, las características de la red se acoplan adecuadamente a las series de tiempo asociadas a los latidos cardiacos, ya que la precisión posee una convergencia adecuada tanto en la fase de entrenamiento como de testeo.

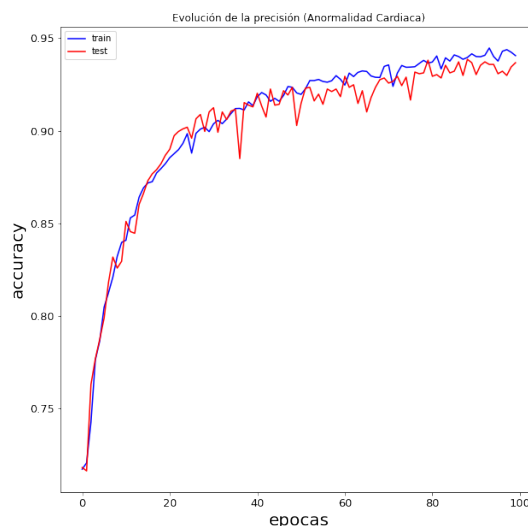


Figura 3.3: Evolución de la precisión del clasificador binario en función de las épocas

Es interesante ver como requiere de tan pocas épocas para lograr una precisión aceptable sin llegar a ser sobrealimentada por la cantidad de parámetros disponibles, lo que demuestra la efectividad de la red para clasificar latidos que poseen afecciones en particular.

Por otro lado, las fluctuaciones mínimas entre las épocas determinan que la convergencia de la red es confiable y limita la incertidumbre, pese a que la base de datos se estima poseer registros de personas con diferentes edades y contextos, lo que es altamente positivo para el reconocimiento general de irregularidades.

Finalmente, la precisión obtenida al clasificar el conjunto de datos de validación y basándose en la probabilidad mas alta obtenida como respuesta del modelo genero una precisión del 93.58 % que es un resultado altamente aceptable que le permite dar cierta flexibilidad al modelo y se estima que bajo un entrenamiento que incurra en más épocas, así como una base de datos aun mayor, los resultados mejorarían considerablemente sin el riesgo progresivo a la sobrealimentación.

Capítulo 4

Redes neuronales para la categorización de irregularidades cardíaca

*El alma nunca piensa sin una imagen
mental*
Aristóteles

4.1. Motivación

Como gran parte de las enfermedades, el primer paso para poder tratarlas adecuadamente es detectarlas con premeditación. En el capítulo anterior se logró realizar una aproximación adecuada para detectar este tipo de irregularidades, generando detecciones altamente precisas de señales no saludables en electrocardiogramas que permitirían apoyar un diagnóstico final.

Sin embargo, parte del proceso de diagnóstico es identificar que tipo de enfermedad aqueja al paciente y bajo que circunstancias su corazón podría tener alguna dificultad o no.

En esta sección el objetivo será complementar la red construida anteriormente implementando un clasificador de perturbaciones cardíaca, específicamente enfocado en las arritmias, de tal manera, la premisa en la que se basará el algoritmo es que, una vez conocida la posible perturbación, desglosar detalladamente que tipo de irregularidad es clasificando las series de tiempo en función a las perturbaciones encontradas.

4.2. Introducción

En el capítulo anterior se comentó la necesidad de una detección temprana de las enfermedades para facilitar el tratamiento adecuado de un paciente, sin embargo, descubrir una anomalía en el corazón es solo la mitad del trabajo.

Una vez reconocida la irregularidad, es necesario realizar una identificación de la afección, ya que esto facilitará encontrar el origen que provocó la enfermedad y proveerá de una correcta medicación, así como una posterior recuperación. Este proceso tiene una complejidad importante ya que existen una innumerable cantidad de enfermedades

y cada día se descubren nuevas, por lo que es usual utilizar algunos síntomas generados por el cuerpo afectado para determinar la causa del problema.

No obstante, la enorme permutación de efectos que una enfermedad puede causar en un individuo no se limita a un conjunto cerrado de anormalidades, sino que efectúan múltiples efectos secundarios que pueden variar en cada persona. A su vez, la existencia de aflicciones previas puede opacar la identificación clara de la anormalidad

Este apartado se centrará de generalizar más los comportamientos cardíacos anteriores clasificando electrocardiogramas en función a las perturbaciones encontradas, para ello en el siguiente apartado se explicarán las modificaciones hechas en la red usada en los capítulos previos y se verá la efectividad que esta tiene para acoplarse a los datos.

4.3. Redes neuronales convolucionales para la clasificación de electrocardiograma

Si bien las características de los electrocardiogramas para la detección de irregularidades no difieren de los electrocardiogramas para la clasificación de anormalidades, el análisis requerido si debe variar al anterior, ya que se necesita de un procesamiento diferente, esto debido a que a diferencia del objetivo anterior, la red de clasificación no solo debe comparar y ajustar un solo modelo ideal de pulsaciones, además debe interpretar las irregularidades del fenómeno y dar un juicio a partir de ello.

Basado en las condiciones del objetivo, parece plantear un enfoque diferente y probablemente una red alternativa solución, sin embargo, las propiedades de resalte de las convoluciones permitirían establecer analogías de forma similar a la de los casos anteriores y adaptar un fenómeno mas específico a las redes resaltando nuevamente las características mas representativas, es decir, los segmentos PR, el complejo QRS y la onda T en la serie de tiempo.

Más específicamente, las enfermedades que se lograron reconocer en este apartado dependen únicamente de las pulsaciones adicionales, faltantes o anormales de los ventrículos, es decir, del comportamiento en el complejo QRS la cual coincide con ser el fenómeno de amplio espectro mas representativo de las señales.

Por lo cual, al ser menos relevante el segmento PR y la onda T las características planteadas para el reconocimiento de anormalidades concuerdan mas con las características del reconocimiento de parámetros en las ondas gravitacionales, las cuales solo basaban su estudio en los tiempos de reposo y activación de mayor o menor amplitud, así como interpretar estos comportamientos a una base guía de referencia.

En este caso particular, la red debería adaptarse a adquirir una interpretación similar a la que se obtenía al entrenar el algoritmo implementando redes neuronales, pero, siendo un estudio mas generalizado de la clasificación binaria de la red de detección.

4.3.1. Especificaciones de la red implementada

Esta variación de la red es una modificación más general del clasificador binario, un clasificador multiclase, por lo que se busca determinar más de dos clases, específicamente 5 clases ya que los datos que se presentaran más adelante solo poseen 5 caracterizaciones, si bien, estas irregularidades solo constituyen una pequeña parte de las posibles aflicciones que se pueden encontrar en la naturaleza, sirve como base para determinar el rendimiento de la red y su capacidad para acoplarse a diferentes estudios y comportamientos.

Nuevamente, con base a la literatura, se utilizó como función de pérdida la entropía, esto con el fin de obtener como resultado un conjunto de valores reales entre cero y 1 que determinan la probabilidad de que la serie pertenezca a esa clase sin ser totalmente

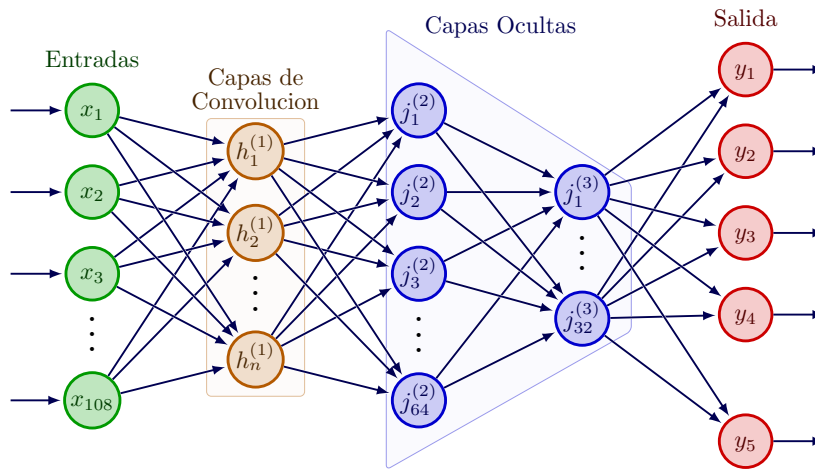


Figura 4.1: Red Neuronal para la clasificación de afecciones cardíaca.

estricto en la clasificación, sino que sea una línea guía con la cual un experto pueda aportar también sus criterios.

Las entradas nuevamente son series de tiempo de 108 elementos, que serán procesadas en 8 filtros de convolución de longitud 64 aplicados cada dos píxeles, activados mediante la función relu, para posteriormente pasar a dos capas densas de 64 y 32 neuronas respectivamente, para ser finalmente activadas mediante una función de activación softmax retornando un vector de 5 categorías correspondientes a la probabilidad de cada clase. La estructura de esta red puede observarse en la figura 4.1.

La métrica de evaluación es la precisión o accuracy y como función de pérdida se utilizara la entropía.

Finalmente, en la siguiente sección se explica el origen de los datos utilizados, la metodología para entrenar la red neuronal y los resultados finales de la red.

4.4. Clasificación de irregularidades

4.4.1. Base de datos

La base de datos es tomada de [19] que contiene 109446 datos seccionados en dos documentos los cuales se clasifican en 5 categorías correspondientes a las categorizaciones de diferentes comportamientos cardíacos los cuales son:

- Normal beat (0): Representa un latido saludable del corazón.
- Supraventricular premature beat (1): Los latidos prematuros supraventriculares representan la activación prematura de las aurículas desde un sitio distinto del nódulo sinusal y pueden originarse en las aurículas o en el nódulo auriculoventricular. Pueden estar asociados con el estrés emocional, la ingesta excesiva de bebidas con cafeína o la administración de agentes simpaticomiméticos (epinefrina, albuterol), incluso con el hipertiroidismo [22].
- Premature ventricular contraction (2): Las contracciones ventriculares prematuras son latidos del corazón adicionales que comienzan en una de las dos cavidades de bombeo inferiores del corazón (ventrículos).

Las contracciones ventriculares prematuras suelen causar pocos síntomas o ninguno. Sin embargo, en ocasiones los afectados logran sentir la sensación de palpitaciones, interrupciones en los latidos, mayor conciencia de los latidos del corazón [14].

- Fusion of ventricular and normal beat (3): Latidos en los que existen movimientos tanto normales como ventriculares.
- Unclassifiable beat (4): Latidos del corazón no reconocidos pero anormales.

La extracción de los datos fue basada en el documento [29] que se enfocaba en la categorización de diferentes arritmias y temas relacionados con afecciones cardíaca

4.4.2. Entrenamiento de la red

Para el entrenamiento se utilizaron los 109446 registros disponibles etiquetados a partir de vectores categóricos de cinco componentes, para el caso de que un registro fuera de un corazón sin alteraciones el vector de etiqueta sería de la forma (1,0,0,0,0) indicando que la probabilidad de que el individuo no tenga alteraciones es de 1, análogamente si la serie de tiempo representa algún latido prematuro supraventricular la etiqueta asociada será de la forma (0,1,0,0,0)

Los datos se distribuyeron en tres conjuntos, 70% del total, lo que equivale a 76612 series de tiempo entrenaran la red, mientras que el resto será distribuido de forma equivalente entre el conjunto de datos de testeo y validación, esto indica que hay 16416 series para cada conjunto.

Como optimizador se utilizó el método Adam, como métrica de evaluación se implementó la precisión o accuracy y como función de pérdida se utilizó la entropía categórica.

En el análisis preliminar del modelo se dejaron interactuar diferentes épocas hasta que se determinó que la convergencia del modelo converge adecuadamente después de las 30 épocas, por lo que se le asignaron 50 épocas de entrenamiento

Finalmente, los resultados del entrenamiento de la red y el rendimiento de esta ante los datos desconocidos durante su ajuste serán mostrados en la siguiente sección.

4.5. Resultados finales

Los resultados del entrenamiento y test indican que en efecto la red logra acoplarse a las irregularidades de las irregularidades cardíaca enfocadas en las arritmias, pese a que la precisión del entrenamiento y test se distancian un poco conforme las épocas incrementan, la convergencia es evidente lo que asegura una correcta categorización de las arritmias en más del 90% de las veces tal y como se muestra en la figura 4.2.

Con respecto a la precisión adquirida en los datos de validación, se obtuvo un 95.84% indicando que, en efecto, la categorización de la red es altamente eficiente y que permite determinar adecuadamente la gran mayoría de las arritmias con gran precisión.

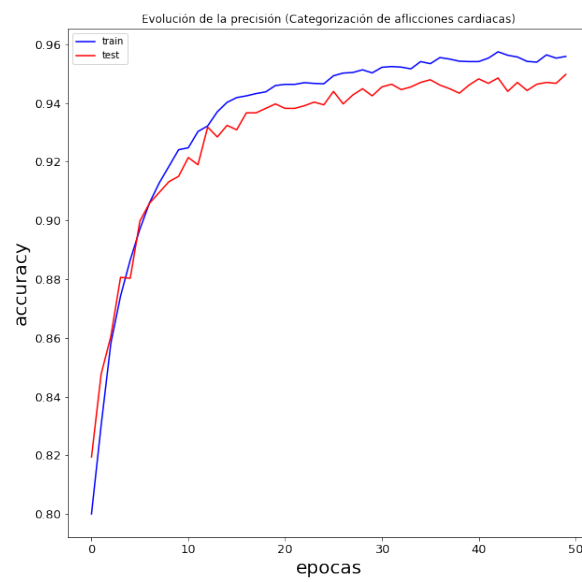


Figura 4.2: Evolución de la precisión del clasificador multiclase en función de las épocas

Capítulo 5

Redes neuronales para identificación de movimientos sísmicos

*El alma nunca piensa sin una imagen
mental*
Aristóteles

5.1. Motivación

Uno de los eventos que han intrigado a la humanidad han sido los movimientos sísmicos, durante los inicios de la historia humana se han registrado el papel que ha tomado los sismos en el desarrollo humano y como estas dificultades han permitido estudios para la generación de conocimiento y ciencia.

Desafortunadamente, las características bajo las cuales estos fenómenos se crean, así como las condiciones bajo las cuales se podrían predecir aun son hipótesis, por lo que establecer modelos de detección autónomos confiables es una herramienta útil para el desarrollo de comunidades con riesgos de desastres naturales provocados por el movimiento de las placas tectónicas.

En esta sección se buscará determinar si un conjunto de señales terráqueas podría o no representar un sismo a futuro, utilizando una base de datos limitada, que, a su vez, permitiría determinar las limitaciones en la red estructurada para el desarrollo de este documento.

5.2. Introducción

Las dinámicas tratadas en los capítulos anteriores son eventos que se originan en base al movimiento de estructuras independientes totalmente definidas, bajo esta premisa, el resultado de la red aplicada ha demostrado buenos resultados para la detección y clasificación, sin embargo, en esta sección se desea explorar la posibilidad de trabajar con fenómenos cuyas interacciones sean dependientes de variables no conocidas lo que repercute en una periodicidad no determinada.

Con base en la anterior premisa, el fenómeno mas apropiado parece ser los movimientos sísmicos generados en la tierra, ya que, pese a todos los estudios empíricos

y teóricos, las variables que influyen en la generación de terremotos no están del todo establecidos [24] y ha sido imposible detectar las condiciones precisas bajo las cuales se producirá un terremoto y en cuanto tiempo lo hará.

Un terremoto se produce cuando dos bloques de la tierra se deslizan repentinamente entre sí. Puede ser inducida por la adición o eliminación de agua superficial o subterránea y por la inyección o eliminación de fluidos debido a la actividad industrial [46].

Dado estos movimientos inestables y la probable destrucción que generan, la localización de terremotos constituye un problema fundamental en el desarrollo de la teoría sísmica y en el establecimiento de comunidades en territorios propensos a estos comportamientos. De forma más detallada, identificar en que momento se podría producir un movimiento es fundamental para ayudar a detectar peligros potenciales en zonas focalizadas, así como caracterizar las actividades del subsuelo y delinear los movimientos de las fallas.

Sin embargo, predecir la aparición de terremotos es una tarea complicada ya que los terremotos no se generan en base un patrón específico y se establecen en base a las irregularidades en los movimientos actuales que podrían o no pertenecer a un movimiento sísmico.

En la actualidad este proceso de detección se podría considerar aún abierto, si bien se han generado muchos avances prometedores en el campo de las detecciones sísmicas [37, 33], dado a que el reconocimiento de movimientos depende de innumerables situaciones como las condiciones del suelo y los patrones sísmicos poco reconocibles, el reconocimiento de patrones no se enfoca únicamente en un solo tipo de metodología, sino que estas son modificadas de acuerdo con el contexto territorial.

Con base en lo anterior, cada detección de sismos requiere de estudios de suelos, así como enfocar el estudio a contextos según la disposición del terreno, presencia de fuente hídricas, entre muchos otros factores ambientales.

Por otro lado, la cantidad de eventos sísmicos que se pueden generar en un territorio podría no ser suficientes para producir un estudio preciso o confiable, ya que los terremotos son eventos que podrían ser efectuados en periodos de tiempo extremadamente largos, lo que evidencia una ausencia de información destacable que frecuentemente imposibilita la implementación de métodos tradicionales teóricos, como algoritmos computacionales más recientes para su predicción.

Una de las técnicas más sofisticadas en el campo de la detección ha sido el Deep learning, que basa su precisión en la calidad de datos con los que se ajusten los modelos, los cuales frecuentemente son redes neuronales con múltiples capas de neuronas, pese a que su principal limitación sea la necesidad de bases de datos lo suficientemente grandes.

Sin embargo, como se ha demostrado en las secciones anteriores, la red neuronal estructurada para el análisis de fenómenos propagados por ondas posee una cantidad muy limitada de neuronas en comparación a los que se implementan en diferentes métodos de Deep learning, lo que garantiza la flexibilidad y generalidad del modelo sin sacrificar su precisión. Además, la convolución generada por el algoritmo teóricamente facilitaría el proceso de reconocimiento de patrones si es que estos existen además de amplificar los picos de amplitud máxima que podrían ser los desencadenantes de irregularidades en las placas tectónicas.

Por lo tanto, en esta sección se explorará el alcance y limitaciones de la red estructurada aplicándola a la predicción de ondas relacionadas con los movimientos sísmicos en los cuales la base de datos tiene una limitada cantidad de información y se determinará la precisión de esta en comparación a otras clasificaciones.

5.3. Las series de tiempo generadas por los movimientos sísmicos

Los terremotos como la gran mayoría de fenómenos naturales son alteraciones en el ambiente que ocurren súbitamente en el interior de la tierra y que produce una dispersión de energía que se propaga a lo largo de la corteza terrestre generando las conocidas perturbaciones en el suelo. Estas propagaciones de energía se generan debido a diferentes causas, una de ellas es la acumulación de rocas en la corteza terrestre que dividen y rompen el subsuelo generando fallas en el suelo para posteriormente generar la liberación de energía cinética propagada sobre la superficie.

Esta liberación de energía posee las características necesarias para considerarse una onda de propagación análoga a la dispersión de energía presentada en la colisión de masas inconmensurables, sin embargo, existen dos principales diferencias con respecto a los movimientos sísmicos.

La primera es el medio de propagación, ya que la energía liberada por el terremoto es dispersada a partir de estructuras sólidas con densidades totalmente estables, a diferencia de las ondas gravitacionales cuyo medio de propagación es el vacío, esto genera perturbaciones importantes en la dispersión de la energía, como fuentes de fricción, irregularidades en la amplitud de las ondas residuales. Aspectos que desestabilizan la frecuencia y el periodo de la onda, dificultando el reconocimiento de la fuente generadora, así como su ubicación.

La segunda diferencia se encuentra en los instrumentos de medición utilizados para evidenciar sus comportamientos, ya que, al cambiar de medio de propagación, las características de sus instrumentos difieren. En particular, los sismógrafos, mecanismos especializados en la medición de sismos, registran todo movimiento terrestre, lo cual puede proveer de diferentes tipos de errores en tiempos y detecciones de manera aleatoria cuyas causas podrían ser ajenas o incluso desconocidas.

El sismógrafo genera mediciones de acuerdo con la magnitud de la onda, esta a su vez puede presentarse en diferentes escalas, pero la más conocida es la magnitud (M_l) de Richter o local [10], la cual fue introducida por Richter en 1935. Se define de manera que un terremoto de magnitud 3 registrado en un sismómetro de torsión Wood-Anderson a una distancia de 100 km escribiría un registro con una excursión máxima de 1 mm.

La sensibilidad de las escalas es adecuada para detectar movimientos que incluso los humanos no pueden percibir, sin embargo, en ocasiones puede ser contraproducente en la predicción de dinámicas, ya que estas pueden generar interacciones que no corresponden a un movimiento sísmico y que permitirían predecirlo como tal, siendo este un problema en la detección temprana de terremotos.

5.4. Redes neuronales convolucionales para la identificación de movimientos sísmicos

De forma particular, las pequeñas variaciones de las ondas en la medición del sismógrafo podrían ser omitidas al aplicar la convolución, no obstante, un porcentaje de estas características podrían ser parte o no del sismo que se pretende predecir, por lo que la estructura fundamental en la que se basaba los anteriores estudios (la convolución) no es del todo confiable para este fenómeno, lo que indica que se debe implementar filtros mucho más leves en este contexto y ajustarlos de manera que se estudie toda la serie de tiempo sin alterar de sobremanera las características originales de la señal.

Actualmente no parece existir un método o estructura especializada que indique la intensidad del filtro que se deba utilizar para destacar adecuadamente las señales pro-

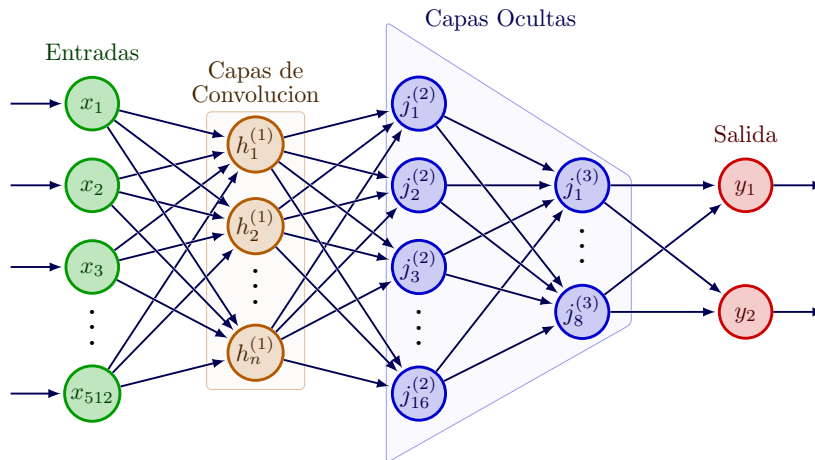


Figura 5.1: Red Neuronal para la predicción de sismos

ducidas por los sismos, por lo cual, el proceso de identificación de intensidad en el filtro debe ser más empírico y es totalmente dependiente a la experimentación ajustada al contexto.

5.4.1. Especificaciones de la red implementada

Esta red es un clasificador binario, ya que se busca generar una predicción en base a dos grados de verdad, si existe la posibilidad de un movimiento sísmico en la serie de tiempo que lo identifique como positivo, de lo contrario será negativo.

Sin embargo, como se especifico en la introducción del capítulo, el comportamiento de estas series de tiempo es muy inestable y dependiente, por lo que no es conveniente generar predicciones certeras, sino que se caractericen a partir de métodos probabilísticos, medidas de incertidumbre que destaquen y apoyen una detección contextual y que posteriormente puedan ser validadas a partir de evidencia empírica.

De manera análoga a los capítulos anteriores, la estructura de la red neuronal que se utilizara será similar, sin embargo, la cantidad de neuronas en este caso variara debido a la limitada cantidad de datos que se posee y la longitud de las series de tiempo no admiten grandes capas de convolución sin una pérdida de información importante, ya que, a diferencia de los casos anteriores, no hay un punto preciso en la serie de tiempo en la que resaltar una sección facilite el proceso de predicción.

Las entradas de esta red son series de tiempo dispuestas en forma vectorial de 512 valores, que serán procesadas a partir de un único filtro de convolución de longitud 2 que será aplicado en intervalos de un píxel ($stride = 1$) activada mediante la función de activación relu, para posteriormente pasar a dos capas densas de 16 y 8 neuronas respectivamente, entre ellas habrán capas de dropout el cual se encargara de omitir aleatoriamente 80% de las neuronas durante el proceso de la red para evitar la sobrealimentación del modelo debido a las limitaciones en la información.

Finalmente, las capas ocultas serán activadas mediante una función de activación softmax retornando un vector de 2 categorías correspondientes a la probabilidad de que exista o no la posibilidad de un movimiento sísmico. La estructura de esta red puede observarse en la figura 5.1.

La función de pérdida utilizada es la entropía y la métrica de evaluación es la precisión o accuracy, que se calcula dividiendo la cantidad de casos favorable o clasificados correctamente, sobre la cantidad de casos totales.

Para este caso se espera tener un rendimiento promedio de un 50 % o menor debido a las condiciones del estudio, las limitaciones de los datos y la falta de contexto, sin embargo, según la calidad de las predicciones este experimento permite detectar bajo que condiciones la red neuronal podría no acoplarse adecuadamente a un fenómeno, siendo este proceso, el que permite detectar el alcance y limitación del modelo estructurado.

Finalmente, en la siguiente sección se explica el origen de los datos utilizados, la metodología para entrenar la red neuronal y los resultados finales de la red.

5.5. Detección de sismos

5.5.1. Base de datos

Los datos se toman del Centro de datos de terremotos del norte de California [5] y cada dato es una lectura promediada durante una hora, con la primera lectura tomada el 1 de diciembre de 1967, hasta la última en 2003.

Cada serie de tiempo es única en su clasificación definiendo cada evento como cualquier lectura de más de 5 en la escala de Richter. Los eventos importantes a menudo son seguidos por réplicas, por lo tanto, se consideran un caso positivo cuando un evento importante no es precedido por otro evento importante durante al menos 512 horas.

Para construir un caso negativo, se considera casos en los que hay una lectura por debajo de 4 (para evitar que los límites entre eventos importantes y no importantes se desdibujen) que está precedida por al menos 20 lecturas en las 512 horas anteriores que no son cero (para evitar casos negativos triviales).

Ninguno de los casos se superpone en el tiempo. De las 86.066 lecturas horarias, se produjeron 368 casos negativos y 93 positivos. Cada serie de tiempo es un vector de 512 elementos.

5.5.2. Entrenamiento de la red

Para el entrenamiento se separaron los 461 registros en dos mitades, 230 para entrenamiento, 230 para test y validación los cuales se dividieron en partes iguales quedando cada grupo con 115 series de tiempo.

De la misma forma que se realizó en el detector de anomalías cardíaca, para este entrenamiento se etiquetó cada serie de tiempo aplicando vectores categóricos de dos componentes, para el caso de que un registro prediga un movimiento sísmico la etiqueta sería de la forma $(0, 1)$ análogamente de no haberlo sería $(1, 0)$

Como optimizador se utilizó el método Adam, como métrica de evaluación se implementó la precisión o accuracy y como función de pérdida se utilizó la entropía binaria, la figura 5.1 ejemplifica la red descrita anteriormente.

Para este caso, debido a la pequeña cantidad de datos, se modificó la red usando solo una capa de convolución, aplicada en un kernel de tamaño 2 utilizado en intervalos de 1 casilla aplicando una función de activación relu.

Se utilizaron dos capas una de 16 neuronas y otra de 8 entre ellas dos capas de dropout del 80 % para evitar la sobrealimentación, así como una función de activación softmax. Finalmente, la cantidad de épocas para el entrenamiento fueron de 80 iteraciones.

Finalmente, los resultados del entrenamiento de la red y el rendimiento de esta ante los datos desconocidos durante su ajuste serán mostrados en la siguiente sección.

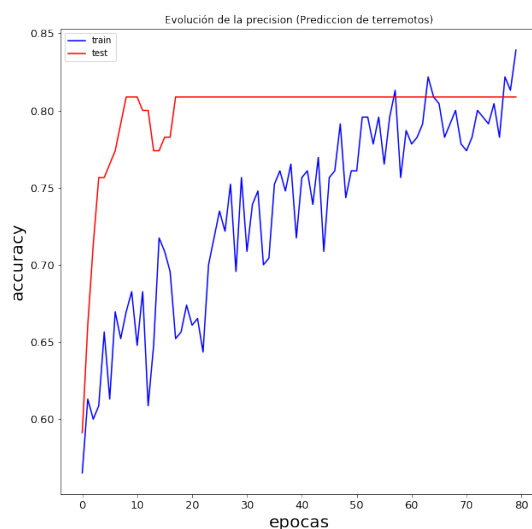


Figura 5.2: Precisión en la predicción sísmica

5.6. Resultados finales

La evolución del proceso de entrenamiento se visualiza en la imagen 5.2, en ella se observa un estancamiento en los datos de testeo mientras que los datos de entrenamiento poseen un incremento progresivo en la precisión, de tal manera, el error en la predicción de los sismos esta encaminada a una detección apropiada, sin embargo, el proceso requiere de una mejor predicción probablemente a causa de la falta de información.

Si bien el resultado es muy inferior a los de los casos anteriores, la evolución del entrenamiento supera las expectativas que se tenían, mejorando incluso la precisión de las categorizaciones mencionadas en [5] del 75.92%, para la predicción en caso de entrenamiento y test se alcanzo un aproximado del 80%

Para la validación, datos que no se dieron a conocer al modelo en su fase de entrenamiento, la precisión fue del 75.86% ajustándose al estándar mencionado en la base de datos y demostrando una buena flexibilidad en la calidad de las predicciones.

En términos generales, el principal problema de la predicción es la limitación de los datos, por lo tanto, se estima que con el doble de datos y aun mejor distribuidos en detecciones y no detecciones, el clasificador tendrá un rendimiento similar al de las secciones anteriores.

Capítulo 6

Redes Neuronales estructuradas bajo el patrón de diseño Facade

*El alma nunca piensa sin una imagen
mental*
Aristóteles

6.1. Introducción

Como ultimo capitulo, en esta sección se desea hacer uso de todo lo establecido en las secciones anteriores integrando cada una de las redes neuronales en un patrón de diseño estructural llamado Facade; [4] Dicho patrón posee la característica de establecer una clase tipo interfaz que permite interactuar con un conjunto de subsistemas para delegar la ejecución de una operación en concreto, sin tener el conocimiento funcional de cada operación en sí. El objetivo de estructurar las funcionalidades de cada red neuronal de esta manera nace de la necesidad de poder clasificar un conjunto de datos de entrada que corresponden a un fenómeno de estudio particular y delegar a una red neuronal secundaria el análisis y la inferencia de los datos esperando obtener los resultados para la cual fue entrenada. Para esto era necesario estructurar una red neuronal convolucional lo suficientemente general que demostrara una efectividad aceptable en múltiples situaciones con el fin de que se adaptara a diferentes contextos manteniendo una confiabilidad medible. Además, identificar en qué condiciones la red podría o no funcionar y estimar el error que este podría generar en la información resultante.

Por ende, en este capitulo se desea bosquejar como el patrón de diseño integra una red neuronal Fachada con otra red neuronal multifuncional que recibe peticiones con datos de entrada previamente clasificados, de esta manera es posible diseñar una aplicación de redes neuronales que trabajan en conjunto unas con otras delegando así diferentes procesos de inferencia, el desarrollo de esta aplicación se hace haciendo uso de clases e interfaces como es sugerido, de esta manera es mas fácil integrar nuevas redes al sistema.

Debido a las limitaciones del estudio, en esta sección solo se utilizó los 4 tipos de series de tiempo implementados en este trabajo, sin embargo, los resultados presentados permiten generalizar estos procesos a fenómenos naturales con los siguientes requisitos:

- Fenómenos cuya propagación pueda determinarse en una serie de tiempo.

- Dinámicas que posean una cantidad de datos aproximadamente mayor a 800 eventos.
- Eventos cuyas dinámicas sean únicamente dependientes de la propagación del fenómeno y no del entorno

Algunos ejemplos que cumplen con estas características podrían ser:

- Diagnóstico y clasificación de síntomas en pacientes con Parkinson.
- Particularidades en el reconocimiento de voz.
- Predicción de precios en la bolsa.
- Fallos en el sistema de un motor de combustión interna.
- Análisis de fluctuaciones cíclicas en sistemas físicos.

6.2. Desarrollo

Muchos de los procesos que se trabajan en la actualidad tienen procesos específicos, métodos independientes que cumplen una función específica. En este trabajo se pudo evidenciar algunos protocolos realizados para generar tareas independientes en base a el análisis de datos temporales, información que frecuentemente nos rodea y que en ocasiones ignoramos.

Todas las señales que constantemente nos rodean son datos de alto interés que podrían indicarnos desde nuestros hábitos poco saludables, algún desastre natural, hasta de eventos astronómicos que puedan ocurrir sin que nos demos cuenta. Si bien podríamos identificar uno que otro comportamiento en nuestro entorno, muchos de ellos son omitidos debido a la limitación de nuestros sentidos.

Por esta razón, el objetivo de este trabajo es establecer las bases de una inteligencia artificial multipropósito que le permita a un ordenador categorizar e interactuar con diferentes entornos, así como desarrollar diferentes tareas en base a los contextos que identifique.

A lo largo de los 4 capítulos anteriores se observó como una red neuronal especializada puede caracterizar adecuadamente comportamientos individuales representados en series de tiempo, por lo que, de la misma forma, estos datos pueden ser seccionados bajo un mismo algoritmo y redirigidos a las estructuras especializadas que se encargan de procesarlas y dar información útil al respecto.

Para ello se utilizará la estructura de la red neuronal usada en los 4 casos y a partir de esta se generará un clasificador que identificará a que fenómeno pertenece cada uno, para posteriormente generar como salida un resultado de acuerdo con la naturaleza del fenómeno. En caso de ser una onda gravitacional retornar las masas de las que se generaron, en caso de ser un ritmo cardiaco identificar si existe una anomalía, en caso de detectar una arritmia identificar de cual se trata y en caso de detectar un movimiento sísmico leve, predecir si este será originario de un temblor de alta magnitud. El esquema 6.1 representa el comportamiento descrito anteriormente.

Para lograr este objetivo basta con generar una última red neuronal que clasifique las series de tiempo en las 4 categorías disponibles, lo cual indica que se deben tomar una muestra equitativa de las 4 bases de datos y analizarlas en una red neuronal.

Las entradas de esta red deben ser series de tiempo deben ser de la misma longitud, lo cual indica que a cada serie se le debe generar un muestreo adecuado que uniformice

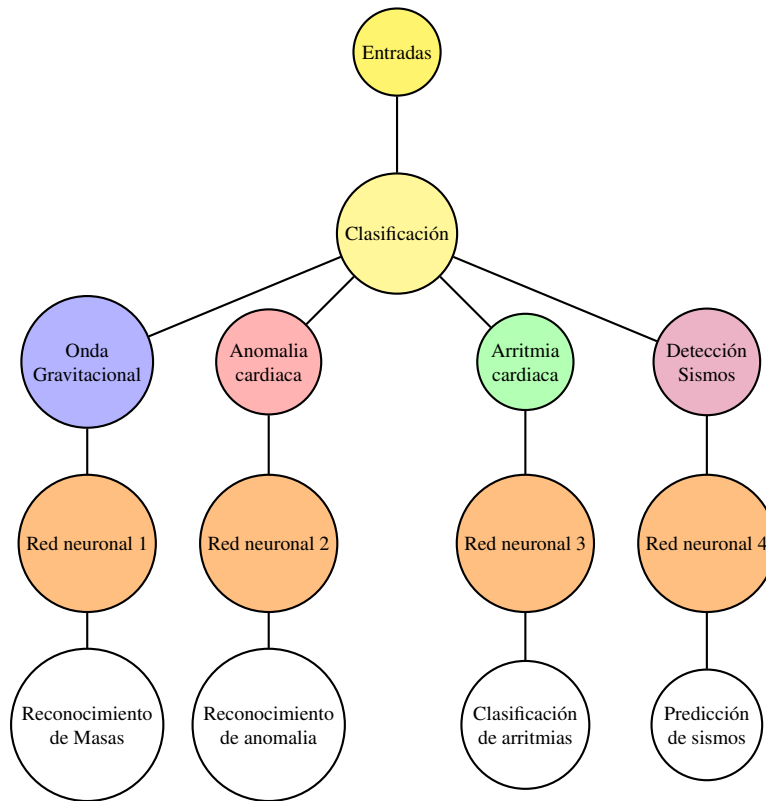


Figura 6.1: Diagrama de distribución del redes neuronales con patrón Facade

la cantidad de datos procesados para posteriormente tomar la señal completa y que una de las cuatro redes ya entrenadas logren realizar el proceso para el cual se diseñaron. De tal manera, las entradas en la red neuronal serán uniformes y podrá detectar a partir de muestras a que fenómeno hace referencia la señal sin tener una aproximación directa del contexto.

Basada en esta premisa, las características del algoritmo implementado deben ser similar a las estructuras con las que se entrenaron los fenómenos, por lo cual, la estructura de red neuronal surge de manera natural ante esta necesidad, ya que ha demostrado gran rendimiento en el reconocimiento, detección y acoplamiento de estos sistemas y de sistemas con características similares variando únicamente pequeños parámetros en su estructuración.

6.2.1. Método de muestreo y entrenamiento

Para calibrar el modelo se tomaron 800 series de tiempo de entrenamiento, 400 para test y 400 para validación. Cada conjunto de datos se dividió equitativamente entre cada grupo de datos, de manera que para el entrenamiento se dispusieron 200 datos pertenecientes a ondas gravitacionales, 200 de anomalías cardíaca, 200 de arritmias y 200 de movimientos sísmicos, de forma análoga para el conjunto de entrenamiento y test.

En caso de las series de tiempo para las ondas gravitacionales se muestrearon tomando

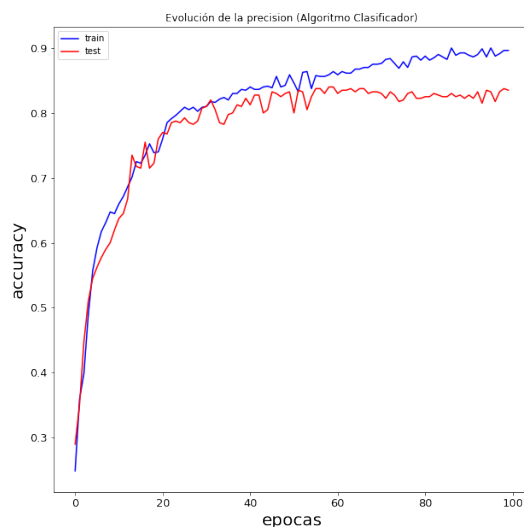


Figura 6.2: Resultados en el entrenamiento del clasificador de dinámicas

cada 20 datos hasta llegar al punto de estabilidad, donde se omitieron los primeros 40 valores correspondientes a la fase inicial del sistema, de tal manera, este muestreo generara una serie de 108 elementos.

Para los datos de anomalías cardíaca y arritmias las características de los sistemas ya están uniformes de acuerdo con la red, ya que estas series ya son series de tiempo de 108 elementos.

Finalmente, para los datos de movimientos sísmicos se muestrearon tomando cada 4 valores en la serie, resultando en series de 108 elementos.

De tal manera, la red es un clasificador multiclase que recibe como parámetro series temporales con longitud 108 el cual representa una señal muestreada de un fenómeno desconocido, inicialmente se procesaran a partir de una capa de 8 filtros de convolución de longitud 64 aplicados cada dos pixeles, activados mediante la función relu, para posteriormente pasar a dos capas densas de 64 y 32 neuronas respectivamente, para ser finalmente activadas mediante una función de activación softmax retornando un vector de 4 categorías.

Esta red será entrenada a partir de 100 iteraciones usando como función de error o perdida la entropía categórica y como métrica la precisión o accuracy.

6.3. Resultados

Los resultados del entrenamiento se muestran en la imagen 6.2 y demuestran la efectividad de la red al detectar adecuadamente los fenómenos mediante las muestras aproximándose casi a un 90% en las 100 primeras fases del entrenamiento, por lo cual, para una sesión de entrenamiento mas abundante se estima generar una aproximación mucho mas precisa.

Por otro lado, las características de la red le permitieron poseer una gran aproximación en pequeñas iteraciones, siendo este el principal objetivo, ya que no solo no requirió de grandes flujos de información, sino que además logro realizar una precisión bastante alta sin sobrealimentarse, uno de los principales problemas con los modelos de Deep

learning.

Con base a estos resultados, la red creada se muestra como una herramienta adecuada para el análisis y estudio de series de tiempo siempre y cuando cumplan las condiciones establecidas al inicio de este capítulo. Sin embargo, existen pequeñas limitaciones con la distribución de categorías y la cantidad de datos.

Pese a esto, la red naturalmente no poseerá los mismos resultados que una estructura específicamente estructurada para el desarrollo de un solo fenómeno, sin embargo, permite establecer una base guía para estudios mas complejos y desarrollos mas generales en el análisis de series de tiempo.

Además, como un proceso adicional, debido a su limitada cantidad de parámetros en comparación a redes mucho mas desarrolladas, al ser optimizada bajo metodologías de teoría de información como la conocida tensorflow lite los procesos aquí presentados podrían ser recreados en pequeños procesadores fácilmente utilizables por cualquier persona.

Capítulo 7

Conclusiones

En este trabajo se estudiaron series de tiempo que poseían picos de energía altamente destacables, como es el caso de las ondas gravitacionales cuya evolución presentaba señales energéticas máximas en el punto de colisión de las masas.

Estas fuentes de información máximas nos permitieron dar estructura a la red neuronal, tomando como hipótesis inicial que, al tomar un análisis enfocado en la amplitud, y la longitud de la onda máxima garantizaría detectar los parámetros de una onda de múltiples espectros. Los resultados mostrados en este estudio indican que esta suposición parece ser correcta, además, puede expandirse a fenómenos de diferentes indoles, incluso si los efectos de la onda no están del todo centrados en el punto máximo de amplitud.

De esto se derivan conceptos mucho más técnicos en el campo del procesamiento y análisis de series de tiempo particularizados a las redes neuronales, como que la cantidad de filtros de convolución óptimos para la identificación de parámetros es inversamente proporcional a la distribución de la información en la serie de tiempo.

Esto se evidencia en la red entrenada con movimientos sísmicos, ya que esta no presentaba un pico particular de energía el cual destacar, sino que la distribución de los movimientos se distribuía uniformemente a lo largo de la señal, por lo que los mejores resultados en el reconocimiento de sismos se obtuvieron al disminuir los filtros de convolución, mientras que los demás fenómenos al poseer una concentración de energía, requirieron de una mayor cantidad de filtros para incrementar la precisión del algoritmo.

Reconocer el efecto de la concentración de información en una señal fue fundamental, ya que generalizar este proceso para diferentes series de tiempo suele ser un proceso determinado a partir de la naturaleza del estudio y los parámetros que se desean obtener de los mismos. Por lo cual, construir una red lo suficientemente flexible que funcione como base guía a diferentes fenómenos es un aporte cada vez mas necesario en la construcción de inteligencias artificiales interactivas que puedan reaccionar a las señales de su entorno.

Otra de las conclusiones más importantes de este estudio en cuanto al uso e implementación de la red fueron las condiciones de uso y precisión, así como las falencias que pueden presentar en diferentes fenómenos que no cumplan dichas condiciones:

- Al contrario de las redes neuronales profundas convencionales, la necesidad estricta de bases de datos densas es disminuida en la arquitectura presentada en este trabajo, esto debido a la forma en como las convoluciones establecen seña-

les fácilmente detectables, lo que en particular no solo flexibilizan el proceso de entrenamiento, sino que además disminuyen la posibilidad de sobrealimentación de la red.

Sin embargo, el hecho de poseer una cantidad considerable de neuronas y capas siempre generara una cantidad mínima de datos que, si bien es reducida, imposibilita a la red analizar fenómenos cuyos eventos sean poco frecuentes, o que los métodos de extracción de información no posibiliten la estructuración de una base de datos con aproximadamente 1000 series de tiempo.

- Por otro lado, se requiere que las señales estudiadas presenten memoria a corto plazo, en otras palabras, que el comportamiento de la señal dependa de las interacciones de la serie de tiempo en un periodo anterior, ya que en fenómenos cuyas interacciones dependen del entorno no alcanza a generar una precisión mayor al 80%, como fue el caso de los movimientos sísmicos. Si bien el entrenamiento de la red muestra cierto acoplamiento a fenómenos dinámicos con su ambiente, la precisión de esta se podría ver más o menos comprometida dependiendo del fenómeno, así como la distribución de la información presentada

Para finalizar, se destaca como el patrón de diseño Facade nos permite poner a trabajar o interactuar redes neurales entre sí para desarrollar un sistema escalable fácilmente replicable y que permita añadir nuevas redes neuronales de diferentes categorías que satisfagan las condiciones anteriormente descritas, siendo este un primer bosquejo de lo que podría ser un modelo interactivo de inteligencia artificial sensible a las señales de su entorno.

Bibliografía

- [1] Oludare Isaac Abiodun, Aman Jantan, Abiodun Esther Omolara, Kemi Victoria Dada, Nachaat AbdElatif Mohamed, and Humaira Arshad, *State-of-the-art in artificial neural network applications: A survey*, nov 2018, p. e00938.
- [2] Maha Alafeef, Mohammad Fraiwan, Hussain Alkhalaf, and Ziad Audat, *Shannon entropy and fuzzy c-means weighting for ai-based diagnosis of vertebral column diseases*, *Journal of Ambient Intelligence and Humanized Computing* **11** (2020), no. 6, 2557–2566.
- [3] Akbar Asgharzadeh-Bonab, Mehdi Chehel Amirani, and Alaeddin Mehri, *Spectral entropy and deep convolutional neural network for ecg beat classification*, *Biocybernetics and Biomedical Engineering* **40** (2020), no. 2, 691–700.
- [4] Kamon Ayeva and Sakis Kasampalis, *Mastering python design patterns: A guide to creating smart, efficient, and reusable software*, Packt Publishing Ltd, 2018.
- [5] A. Bagnall, *Sitio web de clasificación de series temporales*.
- [6] Barry C. Barish and Rainer Weiss, *LIGO and the Detection of Gravitational Waves*, *Physics Today* **52** (1999), no. 10, 44–50.
- [7] Christopher M Bishop et al., *Neural networks for pattern recognition*, Oxford university press, 1995.
- [8] C. M. Biwer, Collin D. Capano, Soumi De, Miriam Cabero, Duncan A. Brown, Alexander H. Nitz, and V. Raymond, *PyCBC inference: a python-based parameter estimation toolkit for compact binary coalescence signals*, *Publications of the Astronomical Society of the Pacific* **131** (2019), no. 996, 17.
- [9] Oscar Blancarte, *Introducción a los patrones de diseño*, México, México DF (2016).
- [10] David M. Boore, *The Richter scale: its development and use for determining earthquake source parameters*, *Tectonophysics* **166** (1989), no. 1-3, 1–14.
- [11] R. Bousseljot, D. Kreiseler, and A. Schnabel, *Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet*, *Biomedizinische Technik* **40** (1995), no. s1, 317–318.
- [12] PA Bromiley, NA Thacker, and E Bouhova-Thacker, *Shannon entropy, renyi entropy, and information*, *Statistics and Inf. Series (2004-004)* **9** (2004), 10.

- [13] Jason Brownlee, *A Gentle Introduction to the Rectified Linear Unit (ReLU)*, Machinelearningmastery.Com (2019), 1–21.
- [14] Mayo clinic, *Contracciones ventriculares prematuras - Síntomas y causas - Mayo Clinic*.
- [15] Personal de Healthwise, *Componentes e intervalos en un electrocardiograma (ECG) | Cigna*.
- [16] Alberto Delgado, *Aplicación de las redes neuronales en medicina*, Revista de la Facultad de Medicina **47** (1999), no. 4, 221–223.
- [17] Sanket Doshi, *Various Optimization Algorithms For Training Neural Network | by Sanket Doshi | Towards Data Science*, 2019.
- [18] Aniruddha Dutta, Tamal Batabyal, Meheli Basu, and Scott T Acton, *An efficient convolutional neural network for coronary heart disease prediction*, Expert Systems with Applications **159** (2020), 113408.
- [19] Shayan Fazeli, *ECG Heartbeat Categorization Dataset | Kaggle*.
- [20] Donal P Finegan, Juner Zhu, Xuning Feng, Matt Keyser, Marcus Ulmefors, Wei Li, Martin Z Bazant, and Samuel J Cooper, *The Application of Data-Driven Methods and Physics-Based Learning for Improving Battery Safety*, feb 2021, pp. 316–329.
- [21] Valeri Frolov and Igor Novikov, *Frolov V.P., Novikov I.D. - Black hole physics.. basic concepts and new developments (Kluwer, 1997)*.
- [22] Ary L. Goldberger, Zachary D. Goldberger, and Alexei Shvilkin, *Supraventricular Arrhythmias, Part I*, Goldberger’s Clinical Electrocardiography, Elsevier, 2018, pp. 130–143.
- [23] Prof. Hans Koch, Dr. Ing. Ralf Bousseljot, and Dr. Ing. Dieter Kreiseler, *PTB Diagnostic ECG Database v1.0.0*, 2004.
- [24] Aitaro Kato and Yehuda Ben-Zion, *The generation of large earthquakes*, Nature Reviews Earth & Environment **2** (2021), no. 1, 26–39.
- [25] Andrzej Królak and Mandar Patil, *The first detection of gravitational waves*, Universe **3** (2017), no. 3, 59.
- [26] Erich L Lehmann and George Casella, *Theory of point estimation*, Springer Science & Business Media, 2006.
- [27] Ivana Martić, Nastia Degiuli, Dubravko Majetić, and Andrea Farkas, *Artificial neural network model for the evaluation of added resistance of container ships in head waves*, Journal of Marine Science and Engineering **9** (2021), no. 8, 826.
- [28] Dr. Yitzar Bañuelos Millán, *Electrocardiograma | Dr. Yitzar Bañuelos Millán*.
- [29] George B Moody and Roger G Mark, *The Impact of the MIT-BIH Arrhythmia Database : History, Lessons Learned, and Its Influence on Current and Future Databases*, IEEE Engineering in Medicine and Biology (2001), no. June, 45–50.

- [30] Antonella Nannavecchia, Francesco Girardi, Pio Raffaele Fina, Michele Scalera, and Giovanni Dimauro, *Personal heart health monitoring based on 1d convolutional neural network*, *Journal of Imaging* **7** (2021), no. 2, 26.
- [31] Federico Nesti, Giulio Rossolini, Saasha Nair, Alessandro Biondi, and Giorgio Buttazzo, *Evaluating the robustness of semantic segmentation for autonomous driving against real-world adversarial patch attacks*, *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2022*, pp. 2280–2289.
- [32] Rich Ormiston, Tri Nguyen, Michael Coughlin, Rana X Adhikari, and Erik Katsavounidis, *Noise reduction in gravitational-wave data via deep learning*, *Physical Review Research* **2** (2020), no. 3, 33066.
- [33] Paul G Richards, Felix Waldhauser, David Schaff, and Won-Young Kim, *The applicability of modern methods of earthquake location*, *Pure and Applied Geophysics* **163** (2006), no. 2-3, 351–372.
- [34] Keith Riles, *Detecting Massive Black Holes via Attometry: Gravitational Wave Astronomy Begins*, *Microscopy and Microanalysis* **23** (2017), no. S1, 4–5.
- [35] Frank Rosenblatt, *The perceptron: a probabilistic model for information storage and organization in the brain.*, *Psychological review* **65** (1958), no. 6, 386.
- [36] Jarrel CY Seah, Jennifer SN Tang, Andy Kitchen, Frank Gaillard, and Andrew F Dixon, *Chest radiographs in congestive heart failure: visualizing neural network learning*, *Radiology* **290** (2019), no. 2, 514–522.
- [37] Muhammad Shakeel, Katsutoshi Itoyama, Kenji Nishida, and Kazuhiro Nakadai, *Detecting earthquakes: a novel deep learning-based approach for effective disaster response*, *Applied Intelligence* **51** (2021), no. 11, 8305–8315.
- [38] Hongyu Shen, Daniel George, Eliu A Huerta, and Zhizhen Zhao, *Denoising gravitational waves with enhanced deep recurrent denoising auto-encoders*, *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2019, pp. 3237–3241.
- [39] Hongyu Shen, EA Huerta, and Zhizhen Zhao, *Deep learning at scale for gravitational wave parameter estimation of binary black hole mergers*, *arXiv preprint arXiv:1903.01998* (2019), 8.
- [40] Noah Snavely, *CS1114 Section 6: Convolution*, Cornell University (2013), 1–6.
- [41] Fung Fung Ting, Yen Jun Tan, and Kok Swee Sim, *Convolutional neural network improvement for breast cancer classification*, *Expert Systems with Applications* **120** (2019), 103–115.
- [42] Vijay Varma, Leo C. Stein, and Davide Gerosa, *The binary black hole explorer: On-the-fly visualizations of precessing binary black holes*, *Classical and Quantum Gravity* **36** (2019), no. 9, 095007.
- [43] Sun-Chong Wang, *Artificial neural network*, *Interdisciplinary computing in java programming*, Springer, 2003, pp. 81–100.

- [44] Wei Wei, Asad Khan, EA Huerta, Xiaobo Huang, and Minyang Tian, *Deep learning ensemble for real-time gravitational wave detection of spinning binary black hole mergers*, *Physics Letters B* **812** (2021), 136029.
- [45] Heming Xia, Lijing Shao, Junjie Zhao, and Zhoujian Cao, *Improved deep learning techniques in gravitational-wave data analysis*, *Physical Review D* **103** (2021), no. 2, 024040.
- [46] Xiong Zhang, Jie Zhang, Congcong Yuan, Sen Liu, Zhibo Chen, and Weiping Li, *Locating induced earthquakes with a network of seismic stations in Oklahoma via a deep learning method*, *Scientific Reports* **10** (2020), no. 1, 1–12.