

III. REDES NEURONALES Y SU APLICACIÓN EN LA INGENIERÍA SÍSMICA

III. REDES NEURONALES ARTIFICIALES Y SU APLICACIÓN EN LA INGENIERÍA SÍSMICA

III.1 CONCEPTOS GENERALES

En sus orígenes las Redes Neuronales Artificiales (RNA) fueron concebidas como un intento para entender y explicar como funciona y opera el cerebro humano. Warren McCulloch un neurofisiólogo y Walter Pitts un matemático fueron los primeros en establecer los fundamentos de la computación neuronal, mediante la analogía que presentaron en 1943 entre la estructura y las funciones de las redes neuronales y los circuitos lógicos digitales (ref 31).

Las neuronas biológicas son células del cerebro cuya función primordial radica en la recolección, procesamiento y emisión de señales eléctricas. Su estructura está integrada por un cuerpo celular o soma, del que se ramifican fibras cortas llamadas dendritas y una sola fibra larga conocida como axón que generalmente llega a medir un centímetro, pero puede alcanzar hasta un metro de longitud (fig III.1). Incluso los axones pueden dar lugar a ramificaciones colaterales junto con la rama principal y por consiguiente la conectividad real de esta célula cerebral puede ser bastante complicada, se sabe que una neurona puede tener 10 y hasta 100,000 interconexiones formando complejas estructuras (ref 32).

Las dendritas reciben las señales de entrada, el soma se encarga de procesarlas mientras que el axón transmite la señal desde el soma hacia otras neuronas. El punto de contacto entre el axón de una célula y la dendrita de otra, recibe el nombre de sinapsis. Se estima que el cerebro humano cuenta con cien mil millones de neuronas, que ejecutan millones de instrucciones indispensables para realizar una vida normal.

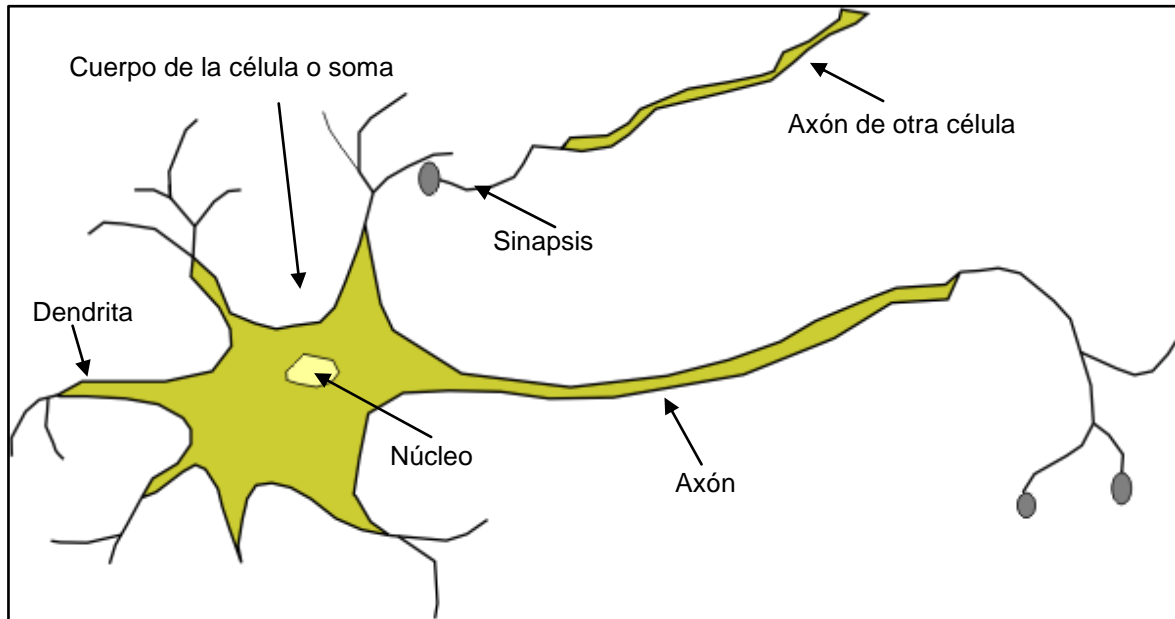


Figura III.1 Neurona biológica (ref 33)

Las RNA's tratan de imitar el comportamiento de la neurona biológica a través de los nodos que son las unidades de procesamiento de estas redes; los nodos igual que las neuronas poseen una gran cantidad de conexiones tanto de entrada de información como de salida (fig III.2). La información recibida se procesa tomando en cuenta los pesos asociados a los nodos W_i , también llamados sinápticos, estos representan la resistencia y el signo de la conexión. Cada nodo tiene un peso asociado, que es la suma de todas las conexiones que llegan al mismo, enseguida de haber sido asignados a todos los nodos, la red evalúa la respuesta que obtiene como salida y ajusta dichos valores para cada nodo.

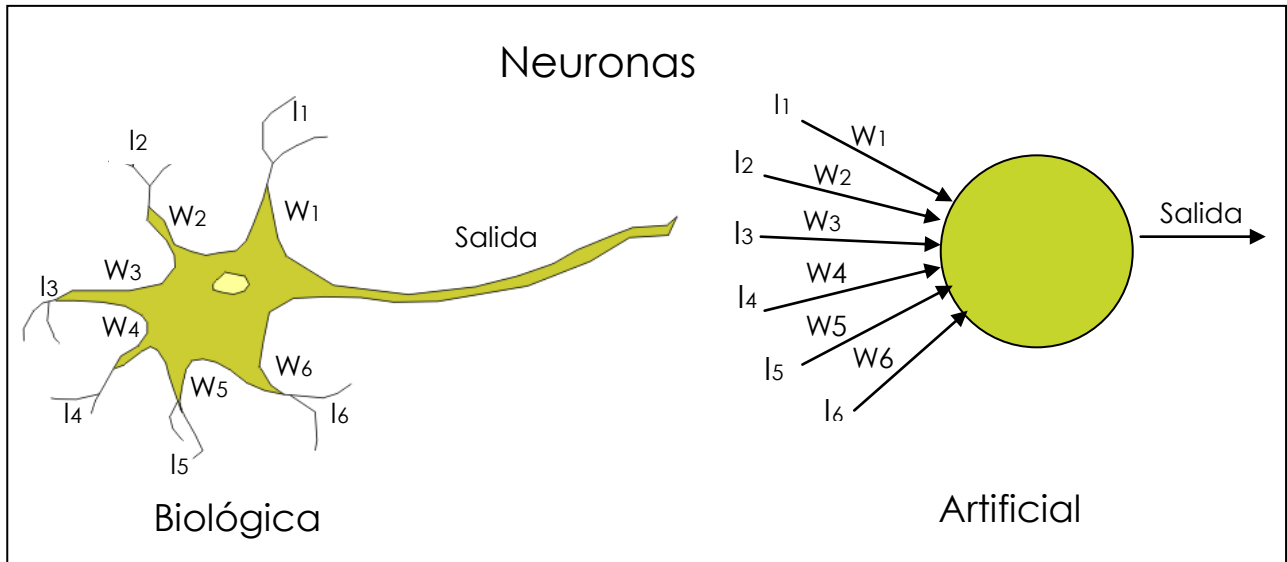


Figura III.2 Comparación entre neuronas biológicas y artificiales

$$I_{tj} = I_1W_1 + I_2W_2 + \dots + I_iW_i + \dots + I_nW_n$$

de donde:

I_{tj} = valor total de entrada al nodo j

I_iW_i = producto punto de la entrada I_i con su peso correspondiente W_i

Las principales características que distinguen a una RNA son: su arquitectura, la dinámica de computación que realiza y el algoritmo de entrenamiento o aprendizaje que utiliza.

Arquitectura o topología: La arquitectura neuronal determina la forma en que se transmite la información, esto es el patrón de conexiones que siguen los nodos en donde las salidas de estos, se convierten en entradas de otros. Los arreglos neuronales

se componen de capas, y de acuerdo al lugar que ocupan en la red se puede clasificar como:

- De entrada: Es la encargada de recibir la información que llega del exterior a la red.
- Ocultas: No mantienen contacto con el entorno exterior y su número varía desde cero hasta un valor grande. Los nodos de las capas ocultas pueden estar interconectados de diversas maneras, dependiendo de la topología de la red.
- De salida: Esta transmite la información hacia el exterior.

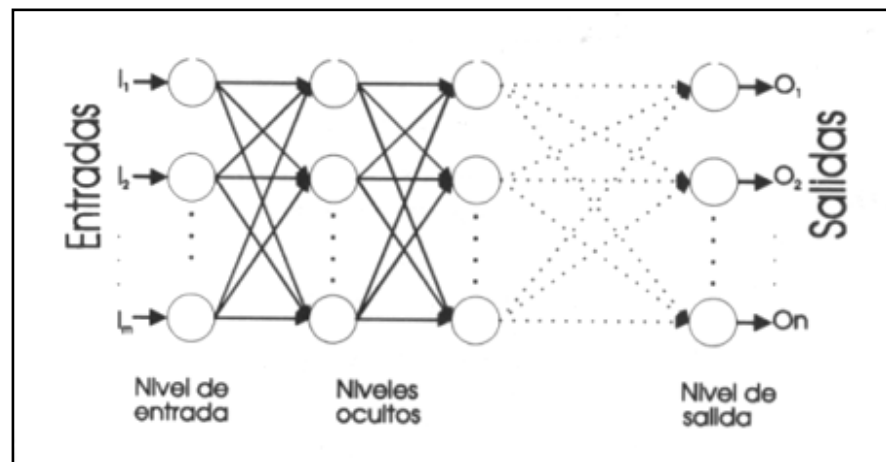


Figura III.3 Capas de una RNA (ref 34)

Los parámetros fundamentales en la arquitectura de una RNA son:

- Número de capas
- Número de nodos por capa
- Tipo de conexiones entre nodos, es decir hacia atrás, laterales o recurrentes (salida de un nodo aplicada hacia su propia entrada).

Existen diversos tipos de topologías uno de los más utilizados es el Perceptrón Multicapa. Dicha arquitectura está compuesta de una capa de entrada cuya única función es transmitir la información de entrenamiento a las capas ocultas, las cuales no tiene contacto directo con el exterior, pero procesan y transmiten la información hacia la capa de salida, la que puede estar integrada de uno o varios nodos. Debido a que el flujo de información se realiza en la dirección que va de la capa de entrada

hasta la de salida, la conexión de los nodos es únicamente con los elementos de la capa siguiente debido a que además está restringida la comunicación lateral con nodos de la misma capa.

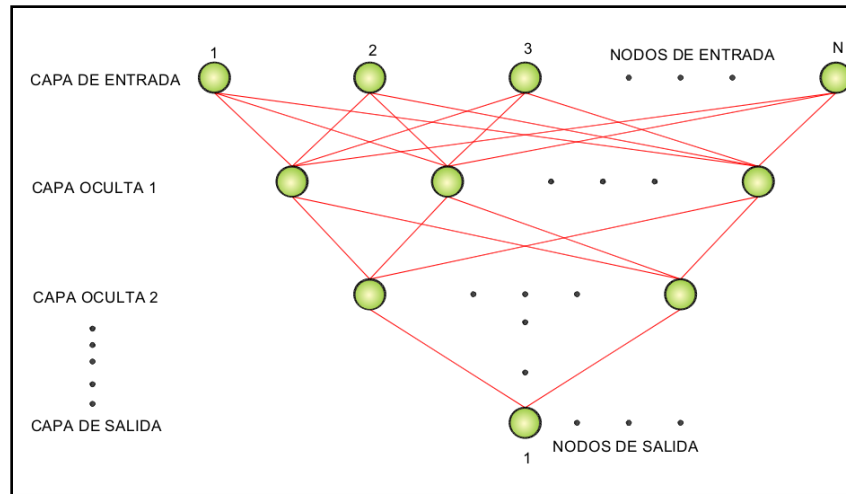


Figura III.4 Arquitectura Perceptron Multicapa (ref 35)

Dinámica de computación: una RNA requiere que sus nodos estén unidos mediante conjuntos de conexiones, asociado a cada uno hay una función de salida o de transferencia, la cual determina si el nodo entra o no, en un estado de activación. Existen diversas funciones de transferencia, entre las más utilizadas se encuentran la escalón, lineal, gaussiana y sigmoideal. La función escalón o umbral (fig III.5) únicamente se utiliza cuando las salidas de la red son binarias (dos posibles valores), se obtiene una respuesta sólo cuando el estado de activación es mayor o igual que cierto valor límite.

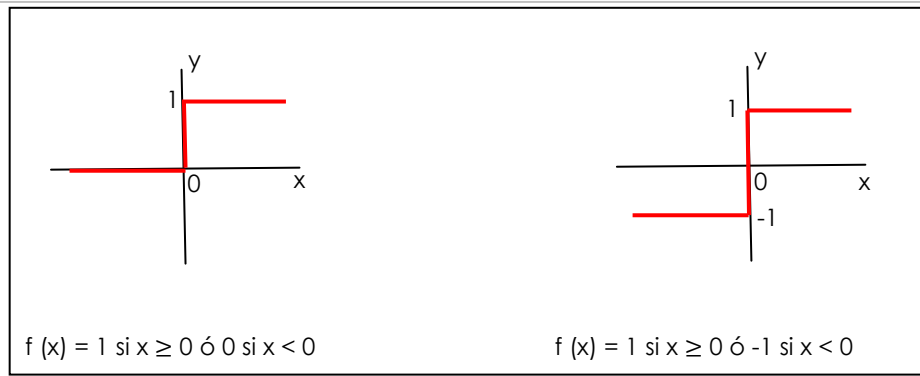


Figura III.5 Función de transferencia escalón

La función de transferencia gaussiana (fig III.6) es más adaptativa que la sigmoideal debido a que el centro (definido por μ) y anchura (en función de σ) de esta se pueden ajustar.

$$FT_{gaussiana} = f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (3.1)$$

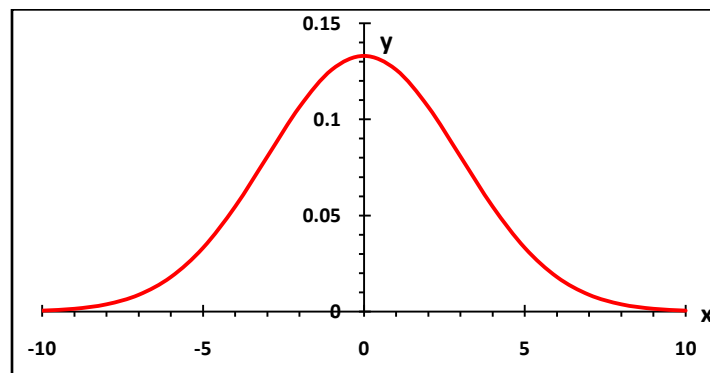


Figura III.6 Función de transferencia gaussiana

La importancia de la función sigmoideal radica en que toma los valores de entrada, que pueden estar entre mas y menos infinito, y restringe la salida a valores entre cero y uno de acuerdo a la siguiente ecuación

$$FT_{sigmoidal} = f(x) = \frac{1}{1+e^{-x}} \quad (3.2)$$

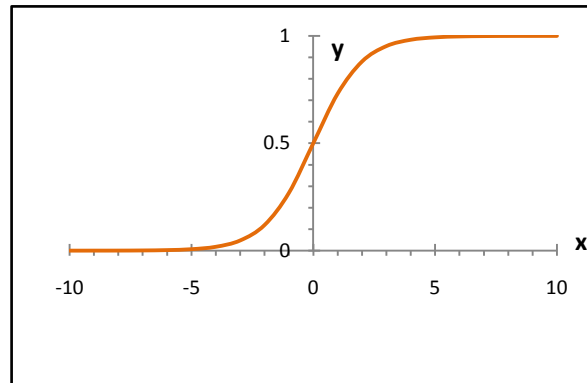


Figura III.7 Función de transferencia sigmoideal

Por otro lado la función de transferencia lineal o identidad (fig III.8) equivale a no aplicar una función de salida, ya que como su nombre lo indica, se trata de una línea recta y por ello no afecta la entrada neta de cada nodo. Es común utilizarla en la obtención de una salida inalterada en la última capa de los arreglos de la RNA.

$$FT_{lineal} = f(x) = x \quad (3.3)$$

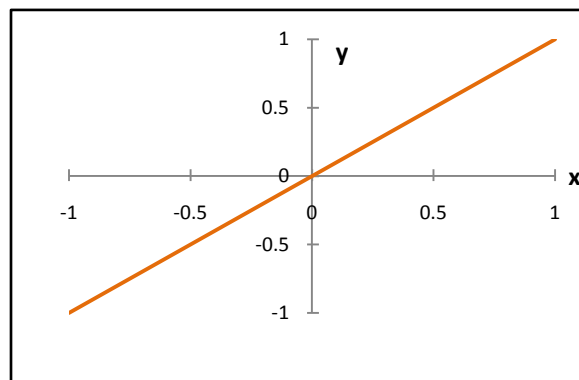


Figura III.8 Función de transferencia lineal

Algoritmo de Entrenamiento o Aprendizaje: de la misma manera en que el funcionamiento de una red depende del número de nodos de los que disponga y la conexión entre ellos, cada modelo cuenta con su propia regla de entrenamiento que representa una parte primordial del proceso de aprendizaje de la RNA así como de su capacidad de generalización. La operación del algoritmo de aprendizaje la integran dos partes fundamentalmente, una fase directa y una inversa o de retroceso. En la primera fase se ingresa a la red un conjunto integrado por pares de entradas y salidas, la información se aplica a la primera capa de entrada y es propagada por las capas ocultas hasta obtener la respuesta real de la red en la capa de salida, estos resultados son comparados con los valores objetivo y se calcula el error, durante esta fase los pesos sinápticos de la red permanecen fijos. En la fase de retroceso el valor del error es transmitido hacia atrás partiendo de la capa de salida, pasando por las capas ocultas hasta alcanzar las capas de entrada, los pesos sinápticos son ajustados de acuerdo con la regla de corrección del error para que en el siguiente ciclo, el valor de las salidas este más próximo a los valores objetivo.

Existen diversos algoritmos de entrenamiento como lo son Propagación Inversa (*Back Propagation*), de Rápida Propagación (*Quick Propagation*), Correlación en Cascada, Gradiente Conjugado, *Levenberg Marquardt* entre otros.

Un problema frecuente que se presenta en los algoritmos de redes multicapa, al buscar minimizar la función de error, es caer en un mínimo local o en algún punto estacionario (fig III.9), con lo cual no se llega a encontrar el mínimo global de la función error. Aunque en algunas aplicaciones no es primordial encontrar el mínimo global, si no encontrar el error mínimo preestablecido, que puede ser una solución aceptable en la práctica (ref 36).

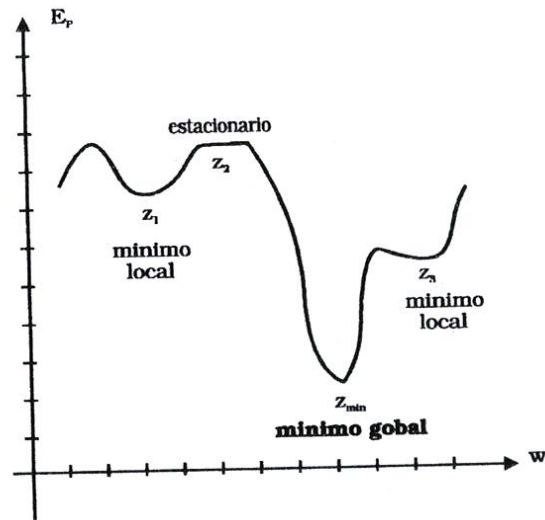


Figura III.9 Mínimos locales y global (ref 37)

III.2 APLICACIONES EN INGENIERÍA SÍSMICA

Las redes neuronales son técnicas de análisis de información desarrolladas para el estudio y manejo, de sistemas mal definidos y difíciles de modelar (ref 38). Son una herramienta para construir modelos “aproximados” contrario a los “precisos” obtenidos con métodos numéricos clásicos. La tecnología computacional que representan las RNA tiene múltiples aplicaciones en diversos campos, algunas aplicaciones son las siguientes:

- Explotación de base de datos
- Optimización de plazas y horarios en líneas de vuelo
- Reconocimiento de caracteres escritos
- Previsión del clima
- Valoración del riesgo de los créditos
- Interpretación de firmas
- Control de producción en líneas de procesos
- Diagnóstico y tratamiento a partir de síntomas y/o de datos analíticos (electrocardiogramas, encefalograma, análisis sanguíneo, etc)
- Predicción de reacciones adversas a los medicamentos
- Clasificación de señales

En la actualidad las RNA's se han aplicado satisfactoriamente en diversos campos de la ingeniería civil, una rama de esta ingeniería es la ingeniería estructural donde, se han utilizado como herramientas para determinar la localización y severidad del daño estructural de manera confiable (ref 39), aunque se han desarrollado distintos métodos para la evaluación de este, como lo son los de inspección visual, emisión acústica, ultrasónicos y dinámicos, todos ellos presentan algunos inconvenientes, el mayor de ellos es que su aplicación no resulta práctica, ya que requieren que la porción de la estructura a inspeccionar sea de fácil acceso, lo cual no siempre sucede, o bien los métodos dinámicos, que se basan en la medición de la respuesta de la estructura (vibración) aún resultan en la mayoría de los casos poco confiables debido a que tienen el inconveniente de no ser lo suficiente precisos para detectar el daño, localizarlo y determinar su severidad, por lo que su uso es un tanto limitada. Una aplicación más de las RNA's en la ingeniería estructural es el desarrollo de modelos para la predicción de la resistencia a la compresión del concreto (ref 40), a partir de variables como la velocidad de pulso ultrasónico y el peso unitario, variables que no son destructivas y si de fácil obtención.

En Geotecnia las redes neuronales han sido una alternativa para modelar el comportamiento de la arena sujeta a esfuerzos de compresión monotónica (ref 41) así como también han representado una nueva opción para la estimación de asentamientos superficiales por tuneleo en los suelos blandos de la ciudad de México (ref 42), este ha sido un tema recurrentemente estudiado debido a la construcción de las obras de infraestructura importantes, como el Drenaje Profundo, el Transporte Colectivo METRO, así como en obras especiales , por ejemplo en la excavación de la Catedral Metropolitana, en este trabajo se comparan los resultados obtenidos mediante los métodos semiempírico, analítico y RNA, obteniéndose la mejor aproximación con esta última.

Dentro de las aplicaciones que la Ingeniería Sísmica ha aportado se encuentran RNA's capaces de modelar la atenuación sísmica, para representar este fenómeno se han desarrollado una gran variedad de leyes de atenuación sin embargo presentan limitaciones, ya que la experiencia en el modelado de los fenómenos sísmicos ha demostrado la no-linealidad, la gama de comportamientos y la imposibilidad para medir algunas variables de estado, dando como resultado que las operaciones de

procesamiento tradicionales resulten poco eficaces en numerosas aproximaciones; en el caso del fenómeno de atenuación sísmica donde las características del comportamiento del sistema se representan por medio de datos de entrada-salida, los arreglos neuronales constituyen una interesante alternativa de modelado (ref 43). Asimismo se cuenta con RNA con capacidad predictiva que pueden estimar espectros de respuesta (ref 44), los cuales son una necesidad para caracterizar los movimientos del terreno bajo acción sísmica y su efecto en las estructuras, de hecho es un medio para representar la respuesta máxima de todos los posibles sistemas de un grado de libertad, para un movimiento particular del terreno, gráficamente el espectro de respuesta es un elemento que integra todos los valores máximos que, como respuesta tiene un sistema de un grado de libertad cuando está expuesto a una excitación en su base. Una aplicación mas es para estimar la duración de la fase intensa (ref 45) cuya interpretación es el tiempo bajo el cual se presenta la parte más relevante del acelerograma; en dicho trabajo se puede apreciar una gran similitud con ecuaciones que describen el fenómeno, dejando en claro la capacidad predictiva que estas redes ofrecen.