

## Modelo De Memoria Usando Red Neuronal De Retropropagación

Ramiro Alejandro Yerovi Vargas  
Departamento de Física  
Escuela Politécnica Nacional del Ecuador  
Ladrón de Guevara E11 - 253. Quito, Ecuador  
yerovi84@hotmail.com

### Resumen

*El trabajo presentado muestra el uso de la Red Neuronal de Retropropagación para el reconocimiento de patrones en tres dimensiones; a través de una simulación computarizada se consiguió reproducir el mecanismo de reconocimiento de objetos. El modelo experimenta aprendizaje sobre objetos en el espacio como un conjunto de datos, además aprende secuencialmente con cada iteración, con lo que se genera un sistema inteligente (inteligencia artificial). Por medio de la simulación se puede recrear, en una escala menor, un proceso que brinda una explicación del funcionamiento de la memoria. En resumen, se utilizó la simulación como una interpretación acerca del funcionamiento de la memoria en base a criterios psicológicos sobre cómo se recuerdan los objetos.*

**Palabras Claves:** Memoria, reconocimiento de patrones, red neuronal, red de retropropagación, inteligencia artificial.

### Abstract

*The present work shows the use of the Back Propagation Neural Network for the recognition of patterns in three dimensions. The object recognition mechanism was modeled through a computer simulation. The model experiment learning over all 3D space as a dataset, also learn sequentially with each iteration, consequently have an intelligent system (artificial intelligence). With the simulation is possible to recreate, on a small scale, a process that provides an explanation of memory. In brief, the simulation is an interpretation about the function of memory based on psychological criteria about how things are remembered.*

**Keywords:** Memory, pattern recognition, neural network, backpropagation network, artificial intelligence.

“El cerebro no es un vaso por llenar, sino una lámpara por encender”

Plutarco

## 1. Introducción

El funcionamiento del cerebro y en especial entender los mecanismos de la memoria han sido siempre temas de gran interés. A pesar de ser un apasionante tema, en sus inicios la investigación sobre cómo funciona el cerebro estaba limitada al poco conocimiento y pocas herramientas de las que se disponía. A pesar de que actualmente se posee mejores métodos y conocimiento, sin embargo todavía no ha sido posible develar completamente su funcionamiento. [1], [2], [6]

Una forma de estudiar el funcionamiento del cerebro es el uso de las redes neuronales que puedan ser simulados por computadora, estos nos dan idea de cómo son algunos procesos que conocemos en el cerebro de forma práctica. La Red Neuronal de Retropropagación, se ha escogido debido a su gran simplicidad en la implementación y a la variedad de aplicaciones que presenta [1], [6], pero permite de una forma minimalista (debido a la incapacidad técnica de modelar todas las neuronas del cerebro) pero bastante apegada a los complejos procesos reales, representar un mecanismo de memoria a través del reconocimiento de patrones. [1], [7], [8], [9]

La memoria es un término muy conocido aunque muy difícil de definir, en su forma más simple podemos decir que la memoria es la forma en que el cerebro almacena cosas, pero en sí lo difícil viene cuando empezamos a detallar cómo es el funcionamiento cerebral que da la capacidad de la memoria, para fines prácticos se va a citar a la memoria perceptual y a la memoria de recuperación, ambos tipos de memoria se basan en procesos cerebrales que se desarrollan en la corteza cerebral y que en el caso de la memoria perceptual es una conexión entre los sentidos y el almacenamiento de la información que los mismos captan, por ejemplo el hecho de tocar algo caliente y almacenar esa sensación y en el caso de la memoria de recuperación es el hecho de recordar cómo se siente la quemadura.[1] Para entender el fenómeno que conocemos como memoria es necesario definir las redes neuronales.

Una red neuronal es un sistema de procesamiento de información que tiene un determinado desempeño que tiene relación con las redes que forman las neuronas en el cerebro. Las redes neuronales han sido desarrolladas como generalizaciones matemáticas de modelos de cognición humana o biología neuronal basada en las siguientes suposiciones:

1. El procesamiento de la información ocurre a un nivel de muchas unidades simples llamadas neuronas.

2. Las señales son pasadas entre neuronas a través de enlaces de conexión.
3. Cada enlace de conexión tiene un peso asociado, el cual, en una red neuronal típica, multiplica la señal transmitida.
4. Cada neurona pasa por la función de activación (que es usualmente no lineal, donde se suman las señales de entrada multiplicada por los pesos y son discriminadas por dicha función) para determinar la señal de salida. [2]

Para describir como están compuestas las redes neuronales es necesario conocer de qué están compuestas, por tanto empezamos con la red neuronal más básica formada por los llamados perceptrones cuyo modelo se aprecia en la figura 1; que son redes neuronales simples que tienen las características citadas anteriormente con la particularidad de que los pesos cambian por cada respuesta del sistema a través de una regla de corrección del error o sea cuando se presenta una respuesta errónea al estímulo que se da en la red [1], [2], [9].

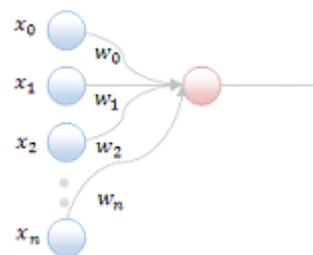


Figura 1. Modelo Perceptrón con sus señales de entrada, pesos y salida

Avanzando en complejidad en cuánto al diseño se describe a la Red Neuronal de Retropropagación como un conjunto de varios perceptrones con la inclusión de capas ocultas, por tal motivo a esta red se le conoce como red multicapa perceptrónica como se ve en la figura 2, porque introduce el modelo de los perceptrones a varios niveles y además corrige la debilidad del perceptrón que es la falta de posibilidades de modelar al operador XOR (conocida comúnmente como “o exclusiva”)[3]; la clave de este modelo es la introducción de las capas ocultas que permiten eliminar las falencias existentes en modelos previos y mejorar la capacidad de la red para aprender.

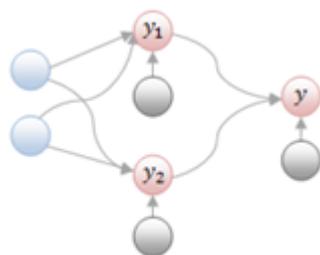


Figura 2. Modelo Multiperceptrón (BPN)

Este trabajo presenta una forma innovadora de entender el mecanismo de la memoria y la percepción sobre objetos en el espacio, de todas las publicaciones y los libros revisados para este trabajo no se ha encontrado un análisis detallado sobre las redes neuronales para el reconocimiento de objetos o patrones en 3D como el presentado sino solo el reconocimiento de letras o figuras simples [1], [2], [3] sin dar mayores detalles de los procesos que realiza nuestro cerebro.

Mediante la red neuronal de retropropagación se busca dar una descripción al mecanismo de funcionamiento de memoria y la percepción desarrollada en objetos en 3D para lo cual se pone a prueba el método introduciendo patrones de objetos tales como un espiral y un vaso para que, mediante entrenamiento de 10000 iteraciones, aprenda la forma de los objetos mencionados. Con lo que se analizan las semejanzas y se identifican las limitaciones del modelo computacional con respecto al mecanismo de memoria presente en nuestros cerebros llegando a conclusiones impresionantes en cuanto a la similitud en cómo se almacenan los recuerdos de objetos y la percepción que se tiene de los mismos.

## 2.- Método

El algoritmo se implementó en un computador personal HP TX2525NR con un procesador AMD de 64 bits con 4GB de RAM, programado en SAS®, que es un programa utilizado por científicos, analistas, consultores en varias ramas como banca, finanzas, etc. [4] Los gráficos presentados en el modelo se realizaron en GNUPLOT. [5] El algoritmo que se implementó en el programa se explica con detalle a continuación.

### 2.1.- Algoritmo de Retropropagación

Los datos utilizados en el modelo son puntos que representan en el espacio a las variables X, Y, Z que luego sirven para representar gráficamente en el espacio a dos estructuras escogidas por el hecho de que tienen forma de objetos familiares a lo que encontramos en el día a día, por tanto se tomó la forma

del espiral de Arquímedes en 3D y la forma de un vaso.

A continuación se detallan las partes de la Red Neuronal de Retropropagación mostrada en la figura 3:

$x_i$  son las entradas del modelo

$v_{ij}$  son los pesos de las entradas enlazadas con la capa oculta

$z_i$  son los valores tomados por la capa oculta

$w_{ij}$  son los pesos de la capa oculta (z) enlazadas con las salidas

$y_i$  son los valores resultantes

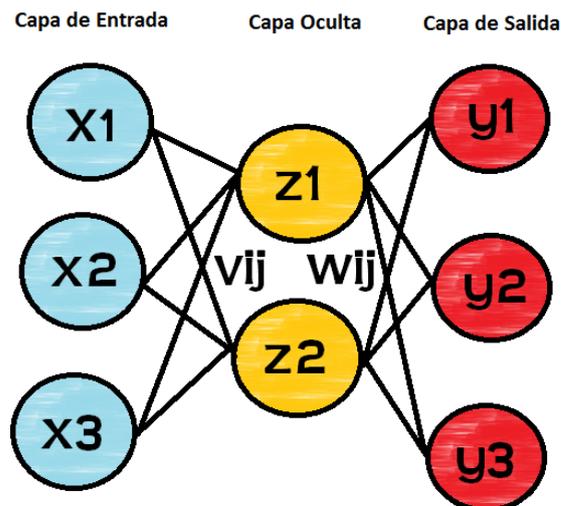


Figura 3. Capas de la Red Neuronal

Además se detalla la notación que se utiliza en el algoritmo:

$Z_{in_j}$  es la suma, de las entradas multiplicadas por los pesos.

$y_{in_j}$  es la suma, de los valores de la capa oculta multiplicados por los pesos.

$f(k)$  es la función de activación aplicada al valor de k.

$\delta_j$  es el valor de la corrección y respectivamente  $\delta_{in_j}$  es la suma, de las correcciones multiplicadas por los pesos.

Por último los valores  $\Delta v$  y  $\Delta w$  son las correcciones en los pesos.

#### Pasos del Algoritmo

Paso 0 Inicializar los pesos asignando valores aleatorios a los mismos.

Paso 1 Mientras la condición para detenerse (la condición la determina el programador en base a la medida del error tolerable) sea falsa haga los pasos del 2 al 9.

Paso 2 Para cada par de entrenamiento haga los pasos del 3 al 8.

- Propagación

Paso 3 Cada neurona de entrada ( $x_i, i = 1, \dots, n$ ) recibe una señal inicial y transmite esa señal

a la capa posterior (capa oculta, puede existir más de una capa oculta).

Paso 4 Cada unidad oculta ( $z_i, i = 1, \dots, p$ ) es sumada una vez que su señal se ha multiplicado con su respectivo peso de la capa de entrada (con los  $v$ ),

$$z\_in_j = v_{0j} + \sum_{i=1}^n x_i v_{ij},$$

Aplicar la función de activación para calcular la señal de salida de esta capa.

$$z_j = f(z\_in_j),$$

Se envía la señal a todas las unidades en la siguiente capa (capa de salida).

Paso 5 Cada unidad de salida ( $y_k, k = 1, \dots, m$ ) se obtiene de la suma pesada de los elementos de la entrada

$$y\_in_k = w_{0k} + \sum_{j=1}^p z_j w_{jk}$$

Y aplicando la respectiva función de activación se obtiene:

$$y_k = f(y\_in_k).$$

#### • Retropropagación de Errores

Paso 6 Cada unidad de salida ( $y_k, k = 1, \dots, m$ ) recibe un patrón objetivo correspondiente al patrón que se tomó como muestra para el entrenamiento, calculando el término de error en la información,

$$\delta_k = (t_k - y_k) f'(y\_in_k),$$

Calculando su término de corrección pesado (utilizado para actualizar  $w_{jk}$  después),

$$\Delta w_{jk} = \alpha \delta_j z_k,$$

Calculando la corrección de los pesos que sirve para posteriormente actualizar el término  $w_{0k}$ ,

$$\Delta w_{0k} = \alpha \delta_k,$$

Y envía  $\delta_k$  a todas las unidades de la capa anterior.

Paso 7 Cada unidad escondida salida ( $z_j, j = 1, \dots, p$ ) sumada su entrada delta del paso anterior,

$$\delta\_in_j = \sum_{k=1}^m \delta_k w_{jk},$$

Multiplicado por la derivada de su función de activación para calcular el error en el término de la información,

$$\delta_j = \delta\_in_j f'(z\_in_j),$$

Calculando la corrección de los pesos que sirve para posteriormente actualizar el término  $v_{ij}$ ,

$$\Delta v_{ij} = \alpha \delta_j x_j,$$

Calculando la corrección de los pesos que sirve para posteriormente actualizar el término  $v_{0j}$ ,

$$\Delta v_{0j} = \alpha \delta_j,$$

Paso 8 Se realiza la actualización de los pesos mediante:

$$w_{jk(\text{nuevo})} = w_{jk(\text{viejo})} + \Delta w_{jk} \quad \text{y}$$

$$v_{ij(\text{nuevo})} = v_{ij(\text{viejo})} + \Delta v_{ij}$$

Paso 9 Condición para detener el algoritmo. [1],[2]

La figura 4 muestra un diagrama esquemático del funcionamiento del algoritmo.

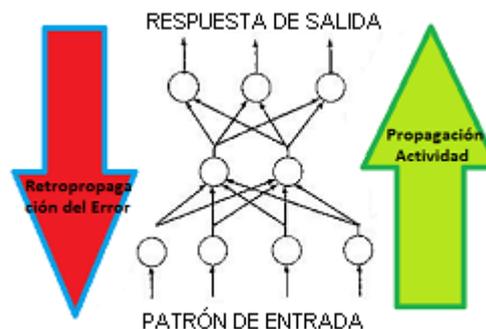


Figura 4. Proceso de la Red de Retropropagación

El modelo diseñado para el presente trabajo tiene una estructura de red 3-3-3 como se muestra en la figura 5.

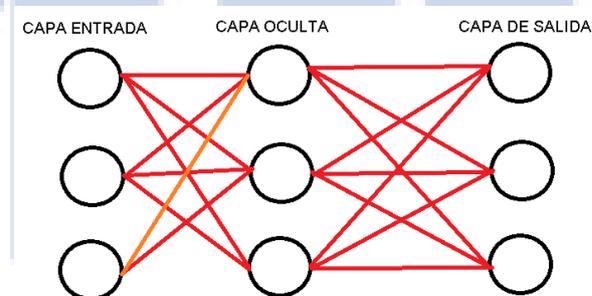


Figura 5. Modelo 3-3-3

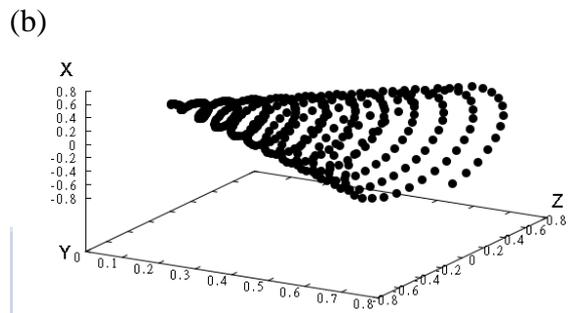
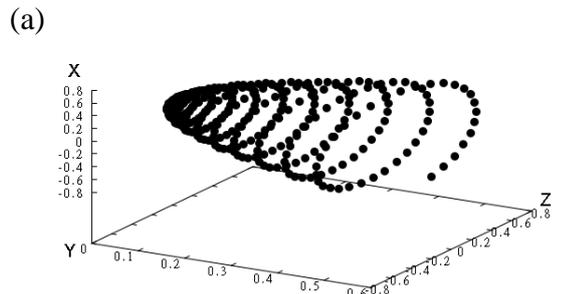
### 3.- Resultados

Para el presente trabajo se tomaron un espiral de Arquímedes en tres dimensiones y una forma de vaso como los patrones a reconocer, dichas figuras se tomaron en cuenta debido a que son comunes y se pueden asociar a objetos manipulados con regularidad en el día a día.

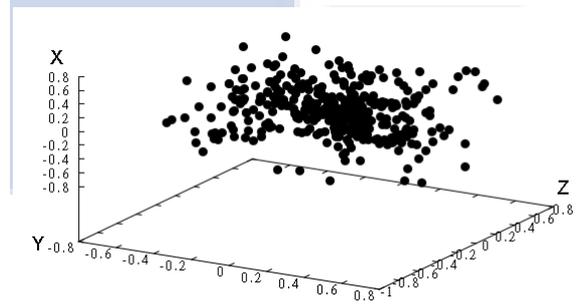
Después de 10000 iteraciones se obtienen resultados de reconocimiento de patrones con menos del 1% en los errores relativos entre el resultado y el patrón de entrada como se puede apreciar en las figuras 7-9.

El algoritmo empieza con un conjunto aleatorio de puntos que es el mismo para ambos procesos y un conjunto aleatorio de pesos como se ve en la figura 7, que empiezan optimizados para que el algoritmo de aprendizaje sea mejor y más rápido. Conforme avanzan las iteraciones del algoritmo de retropropagación, se ve que cuando se llega a 100 iteraciones se obtienen resultados sorprendentes ya que la forma de las figuras es asimilada casi en un 60% como se puede apreciar en la figura 8.

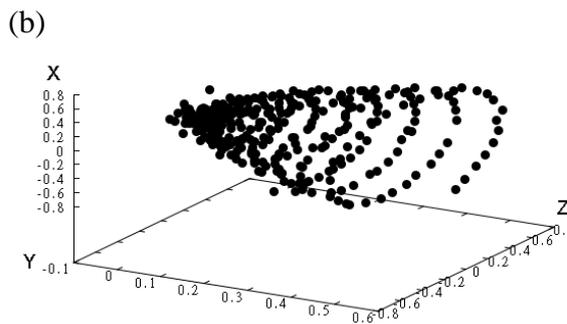
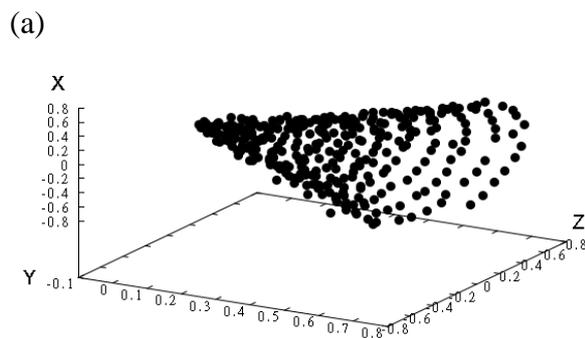
Después de 10000 iteraciones se consigue un resultado igual al conjunto patrón de entrada como se ve en la figura 9.



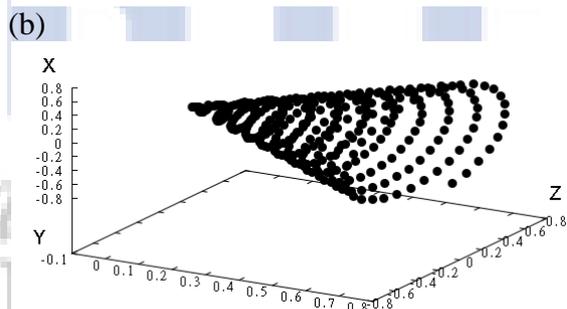
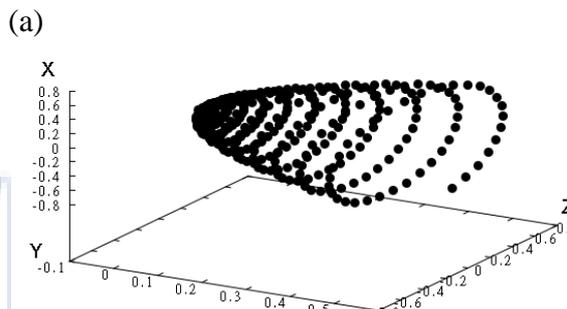
**Figura 6. Patrón de Entrada**  
(a) Vaso  
(b) Espiral de Arquímedes



**Figura 7. Conjunto Inicial de Puntos Aleatorios para ambos procesos**



**Figura 8. Conjunto en la Iteración 100**  
(a) Vaso  
(b) Espiral de Arquímedes



**Figura 9. Resultado Conjunto Final**  
(a) Vaso  
(b) Espiral de Arquímedes

#### 4.- Discusión

En primera aproximación existen varias similitudes entre la memoria convencional y el modelo de retropropagación; vale la pena aclarar que este modelo se lo aplica a la memoria que nos faculta con el reconocimiento de patrones. Por ejemplo, si nunca hemos visto un objeto y lo percibimos instantáneamente tenemos una idea vaga del mismo, pero conforme vamos interactuando con él, lo vemos y nos acostumbramos al mismo, sabemos que no cambia bruscamente su forma aunque cambien imperceptiblemente algunas de sus propiedades físicas a nivel molecular u atómico.

Lo que en realidad va cambiando con respecto del objeto está en nuestro cerebro y es lo que conocemos como percepción [2], [6], [9]; hasta que en un

momento dado y al cabo de cierto número de interacciones somos capaces de reconocer el objeto con gran precisión. En analogía con el modelo planteado se tiene lo siguiente:

Se toma un patrón de puntos fijos que vendrían a representar el objeto que va a ser reconocido, a continuación se empieza con un conjunto de puntos colocados en forma aleatoria en el espacio y la interacción con el objeto que mejora nuestra percepción vendría a ser el ajuste de pesos que realiza el modelo que conforme van pasando las iteraciones se van reajustando para llegar al reconocimiento del patrón.

Debido a que nuestro cerebro tiene  $10^6$  neuronas y el modelo utilizado en la simulación trabaja con 3 neuronas de entrada, 3 neuronas escondidas y 3 neuronas de salida que dan un total de 9 neuronas, comparando 9 neuronas a un total de  $10^6$  el modelo planteado simplemente parecería fuera de lugar e incluso una simple y burda aproximación de lo que en realidad es la memoria, sin embargo la memoria trabaja de una manera sorprendente y se asemeja al modelo planteado en este trabajo y se tienen resultados experimentales que validan y justifican el modelo de redes neuronales de retropropagación como modelos de memoria y de reconocimiento de patrones ya que en psicología algunos criterios que identifica a la memoria como tal son: reconocer objetos, asociarlos con otros objetos y posteriormente reproducirlos [1], [8].

En un estudio realizado en África e Inglaterra se compararon dos colonias de personas totalmente diferentes, en un grupo ubicaron a estudiantes escolares con preparación en Geometría en un nivel aceptable y en el otro a un grupo de nómadas africanos llamados Himba que no han tenido instrucción escolar en toda su vida. Lo sorprendente es que los Himba, pudieron asociar las figuras geométricas de la misma forma en la que lo hicieron los sujetos de prueba de la contraparte occidental que asisten a la escuela [6].

Aplicando el modelo de redes neuronales de retropropagación se puede deducir que el cerebro posee una forma innata de reconocimiento de patrones.

## 5. Referencias

- [1] Michael A. Arbib. The Handbook of brain theory and neural networks, 2002, MIT Press.
- [2] Fausett Laurene. Fundamentals of Neural Networks: Architectures, Algorithms and Applications, 1993, Pearson.
- [3] Antonino Raffone, Jaap M. J. Murre & Gezinus Wolters 2003. NMDA Synapses Can Bias Competition Between Object Representations and Mediate Attentional Selection. Behavioral and Brain Sciences 26 (1):100-101.
- [4] SAS INSTITUT, <http://www.sas.com/>.
- [5] GNUPLOT. Graphics, 1996, vol. 14, pp. 33-38.
- [6] Wickelgren, E. & Bingham, G.P. Trajectory forms as information for visual event recognition: 3D perspectives on path shape and speed profile. Perception & Psychophysics.
- [7] Chessa, A.G., & Murre, J.M.J. 2006. Modelling memory processes and Internet response times: Weibull or power-law? Physica A, 366, 539-551.
- [8] Meeter, M., Talamini, L.M., Schmitt, J.A.J., & Riedel, W.J. 2006. Effects of 5-HT on memory and the hippocampus: Model and data. Neuropsychopharmacology, 31 (4), 712-720.
- [9] Meyering, A. Ritter, H. Learning 3D-shape perception with local linear maps Dept. of Inf. Sci., Bielefeld Univ 1992.