

Introducción a las Redes Neuronales

Luis Fernando Espino Barrios
Instituto Tecnológico de Costa Rica
luisespino@yahoo.com
Diciembre 2009

Resumen: El presente artículo es una introducción a las redes neuronales desde una perspectiva comprensiva, tratando temas como su definición, historia, modelos, funciones, arquitectura, procedimientos, algoritmos, diferentes aplicaciones y un breve estado del arte respecto a herramientas de software relacionadas con las redes neuronales. El principal objetivo es mostrar conceptualmente términos relacionados y la manera como estos son asociados para lograr un objetivo en común. Asimismo, se muestran diferentes puntos de vista en la analogía de las redes neuronales con el sistema nervioso humano y cómo funciona el procesamiento distribuido.

Palabras clave: Funciones de activación, modelos neuronales, *perceptron*, redes de Hopfield, redes neuronales, retro-propagación.

1. Introducción

Las redes neuronales [1] están compuestas de nodos o unidades conectadas por enlaces dirigidos. Un enlace de una unidad a otra sirve para propagar la activación, cada enlace tiene un peso asociado que determina la fuerza y señal de la conexión.

También las redes neuronales [2] permiten el desarrollo de expresiones cuantitativas sin comprometer la complejidad conocida de un problema.

En el resto del artículo se tratarán a detalle las redes neuronales, la arquitectura, los procedimientos, los algoritmos, la aplicación y las principales herramientas de software utilizadas.

2. Redes neuronales

2.1. Definición

Una neurona es una célula del cerebro que tiene como función principal la colección, procesamiento y diseminación de señales eléctricas, análogamente desarrollado en la inteligencia artificial para crear las redes neuronales.

Las redes neuronales [3] son procesadores distribuidos paralelos, hechos para ser unidades de procesamiento simple, que almacenan conocimiento experimental y lo hacen disponible para su utilización. Son llamadas también neuro-computadoras, redes conexionistas, procesadores distribuidos paralelos.

El concepto conexionismo [4] se refiere a un conjunto de modelos que son inspirados en la arquitectura neuronal, un modelo de este tipo es una estructura que está compuesta por un conjunto grande de unidades llamadas neuronas formales, cada unidad está conectada a un subconjunto de otras unidades y cada conexión se representa por medio de un peso

2.2. Modelo de una neurona

Una neurona [3] es una unidad de procesamiento de información, que es fundamental en la operación de las redes neuronales, hay tres elementos básicos en el modelo neuronal:

- Un conjunto de sinapsis o enlaces de conexión, cada uno caracterizado por su peso o fuerza. A detalle, una señal de la entrada de una sinapsis conectada a una neurona es multiplicada por el peso sináptico. Una diferencia de una sinapsis cerebral es que el peso sináptico en una neurona artificial puede encontrarse en un rango que incluye tanto a valores negativos como positivos.
- Un sumador que realiza la suma de las señales de entrada, pesadas por sus respectivas sinapsis de la neurona, las operaciones descritas constituyen un *linear combiner*.
- Una función de activación para limitar la amplitud de la salida de la neurona, es referida también como una función de aplanamiento, en que aplanan los límites o el rango de amplitud permisible de una señal de salida a un valor finito.

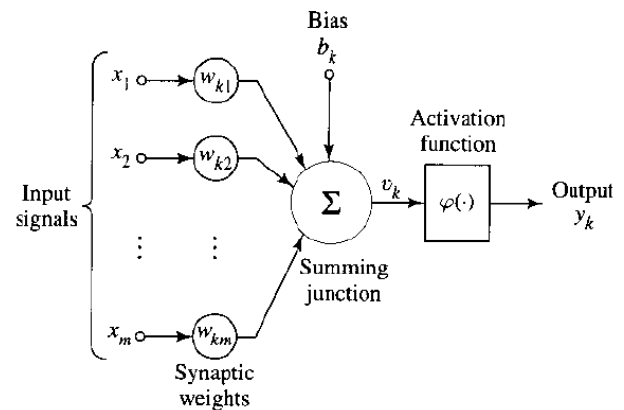


Figura 1: Modelo no lineal de una neurona [3] (p.11)

Se verá más adelante en la sección de historia, que este modelo está basado en el sistema nervioso hipotético llamado *perceptron*.

2.3. Tipos de funciones de activación

Las funciones de activación definen la salida de una neurona en términos del campo inducido.

En [3] se identifican tres tipos básicos de funciones de activación:

- Función de umbral: Comúnmente conocida como función Heaviside, en donde la salida de una neurona recibe tratamiento con una función de umbral previa. Por ejemplo, una función de umbral sería asignar un valor de 1 si el valor es mayor o igual que 0, o de lo contrario tendrá un valor de 0.
- Función de segmentos lineales: En donde el factor de amplificación está dentro de una región lineal de operación asumida por la unidad.

Por ejemplo si hay un rango de valor de cierto valor mínimo a cierto valor máximo recibe el mismo valor de entrada, adicionalmente de las regiones de umbral para devolver valores 1 o 0.

- Función logística: Tiene la particularidad de tener un forma de S, es una función común de activación utilizada en la construcción de redes neuronales, describe una función incremental que exhibe un balance entre el comportamiento lineal y no lineal

2.4. Redes neurales como grafos dirigidos

Tal como se ha visto en el modelo neuronal se describen varios elementos, que pueden simplificarse a través de grafos de flujos de señales sin sacrificar la funcionalidad del modelo, esto con base a un conjunto bien definido de reglas. Un grafo de flujo de señales es una red de enlaces dirigidos que se interconectan en ciertos puntos llamados nodos.

El conjunto de reglas está definido en [3] y son:

- Las señales fluyen a lo largo del enlace solo si la dirección está definida por la flecha del enlace, considerando enlaces sinápticos (relación lineal de entrada y salida) y de activación (relación no lineal de entrada y salida).
- Una señal de nodo es igual a la suma algebraica de todas las señales entrantes en el nodo pertinente a través de los enlaces entrantes.

- La señal de un nodo es transmitida a cada enlace de salida originándose desde el nodo, con la transmisión que inicia completamente de forma independiente de las funciones de transferencia de los enlaces de salida.

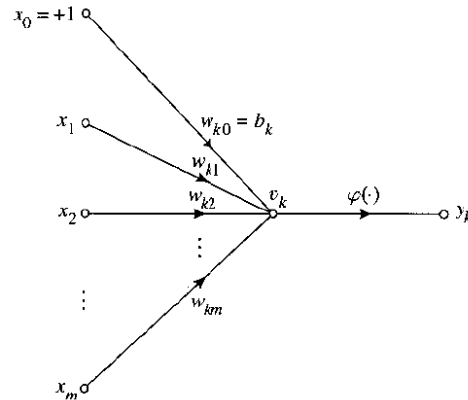


Figura 2: Representación de una red neuronal como grafo dirigido [3] (p.17)

2.5. Definición matemática de una red neuronal

Basado en la definición de [3] se puede definir una red neuronal como un grafo dirigido que consiste de nodos con interconexiones sinápticas y enlaces de activación, y se caracteriza por las siguientes propiedades:

- Cada neurona es representada por un conjunto de enlaces sinápticos lineales, un sesgo externo aplicado y posiblemente un enlace de activación no lineal, el sesgo es representado por un enlace sináptico conectado a una entrada fija de +1.

- Los enlaces sinápticos de una neurona cargan sus respectivas señales de entrada.
- La suma ponderada de las señales de entrada define el campo local inducido de una neurona en cuestión.
- El enlace de activación aplanar el campo local inducido de una neurona para producir una salida.

2.6. Tipos de redes neuronales

En [5] se menciona la clasificación de redes neuronales a través de cinco categorías:

- Modelo a nivel neuronal: Este modelo se concentra en las propiedades de las neuronas individuales, incluyendo el tipo de cálculo de suma ponderada y el tipo de función.
- Modelo agregado: Es la construcción de muchas neuronas, cada neurona tiene propiedades especificadas por el modelo de nivel neuronal, además, se exhiben las capacidades para alcanzar tareas simples.
- Modelo a nivel de red: Está compuesto por un gran número de neuronas individuales, se enfatiza en el comportamiento colectivo y capacidades de la red.
- Modelos a nivel de sistema nervioso: Usualmente construido con dos o más modelos a nivel de red, cada una con diferentes propiedades, la combinación de redes es capaz de procesamiento abstracto y complejo.

- Modelos a nivel de operación mental: Es el modelo de más alto nivel de abstracción, son diseñados para simular y estudiar el proceso cognitivo de creatividad y solución de problemas.

2.7. Beneficios

En [3] se mencionan ciertas capacidades útiles en la utilización de redes neuronales, entre las cuales están:

- Linealidad: Una neurona artificial puede ser lineal o no lineal, siendo una propiedad importante en el mecanismo físico subyacente responsable de la generación de una señal de entrada.
- Mapeo de entrada y salida: Utilizando el paradigma de aprendizaje supervisado, cada ejemplo consiste en una señal de entrada única y su correspondiente respuesta deseada.
- Adaptabilidad: Las redes neuronales tienen la capacidad de adaptar sus pesos sinápticos para cambiar el ambiente a su alrededor.
- Respuesta con evidencia: En el contexto de clasificación de patrones, una red neuronal puede ser designada para proveer información no solo acerca de un patrón particular a seleccionar sino también acerca de la confianza en la decisión tomada.
- Información contextual: El conocimiento es representado por la estructura y el estado de activación de la red neuronal.

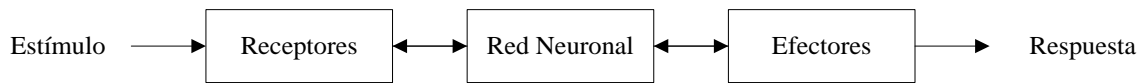


Figura 3: Representación en diagrama de bloques del sistema nervioso [3] (p.6)

- Tolerancia a fallos: Una red neuronal implementada en hardware tiene el potencial de ser inherente a tolerancia a fallos o es capaz de una computación robusta.
- Facilidad de implementación de VLSI: La naturaleza paralela de las redes neuronales hacen que ciertas tareas se ejecuten de manera rápida, de igual forma se hace cuando se implementa con *very-large-scale-integrated* (VLSI), siendo el principal beneficio el capturar el comportamiento complejo de una manera jerárquica.
- Uniformidad de análisis y diseño: Por ser las redes neuronales procesadores universales, se utiliza la misma notación para todo dominio.
- Analogía neurobiológica: El diseño de las redes neuronales es motivado por la analogía con el cerebro, que es la prueba viviente que el procesamiento paralelo con tolerancia a fallos no es solo físicamente posible sino que también es rápido y robusto.

Uno de los principales ventajas se menciona en [2], es que tienen la capacidad de indicar el ruido asociado con la salida cuando se trata con problemas complicados.

3. Historia

3.1. Origen de las redes neuronales

En [3] se menciona la analogía de las redes neuronales con el sistema nervioso humano, este último ha sido la inspiración de dichas redes.

Lo importante de destacar es que existe un estímulo y el sistema devuelve una respuesta, este proceso pasa a través de los receptores y su salida se realiza a través de los efectores, en medio, como parte central del sistema está el cerebro representado por una red neuronal, tal como se observa en la Figura 3.

3.2. Redes de Hopfield

Es un modelo basado en aspectos de neurobiología [6], adaptado a circuitos integrados, las propiedades colectivas de este modelo producen un contenido de memoria direccional, que correctamente cede memoria completa de cualquier sub parte de tamaño suficiente. El algoritmo para la evolución del tiempo del estado del sistema se basa en el procesamiento paralelo asíncrono.

Fue debido a la introducción de las redes neuronales como una teoría de memoria, interesando las siguientes características [7]: representación distribuida, control asíncrono distribuido, memoria de contenido direccional y tolerancia a fallos.

El funcionamiento de la red es como sigue: una unidad aleatoria es seleccionada, si cualquier de los vecinos está activo, la unidad computa la suma ponderada en las conexiones en donde tiene vecinos activos. Si la suma es positiva, la unidad se vuelve activa, de otra manera sigue siendo inactiva. Otra unidad aleatoria es seleccionada y el proceso se repite hasta que la red se encuentra en un estado estable, esto proceso es llamado “relajación paralela”.

La máquina de Boltzmann [7] es una variación de la idea de las redes de Hopfield, se trata de recordar que pares de unidades en un red de Hopfield están conectadas por pesos simétricos.

Las unidades actualizan sus estados de manera asincrónica, buscando en sus conexiones locales de otras unidades.

3.3. Perceptrons

Un *perceptron* es una máquina o sistema nervioso hipotético desde la perspectiva empirista [8]. Es diseñado para ilustrar algunas propiedades fundamentales de los sistemas inteligentes en general, haciendo evidente la analogía con los sistemas biológicos.

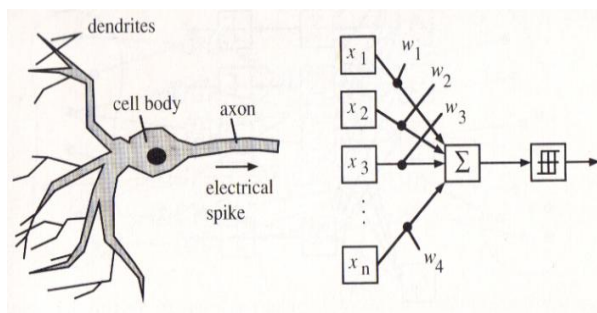


Figura 4: Analogía entre una neurona y un *perceptron* [7] (p.493)

Este fue uno de los primeros modelos de redes neuronales, un *perceptron* modela una neurona que toma una suma ponderada de sus entradas y envía un salida de 1 si la suma es mayor que cierto umbral ajustable [7], a diferencia de las redes de Hopfield las conexiones son unidireccionales.

3.4. Desarrollo posterior

A principios de los 1980's [4], las redes neuronales reaparecieron para tomar dentro de cuentas sub simbólicas tareas como el reconocimiento de patrones, considerándose las redes neuronales como una parte del campo de la inteligencia artificial, desde el punto de vista conexionista en oposición de puntos de vista simbólico.

Al principio estos enfoques se realizaron sin dimensiones temporales, pero más adelante se agregó esta funcionalidad.

El tiempo se logra integrar en las redes neuronales a través de niveles, debido a que el tiempo hay que transformarlo a unidades espaciales para interpretarlo semánticamente [4], las dos maneras de proceso de tiempo fueron a través de mecanismos externos e internos. Con mecanismos internos pueden ser de manera explícita en su arquitectura o de manera implícita en redes recurrentes. De manera explícita puede ser a nivel de red o nivel de neurona. A nivel de neurona puede ser de tres tipos: por sesgos, algebraico o integrado.

La reaparición de las redes neuronales en parte se debió [1] a estudios realizados por el físico John Hopfield 1982, donde utilizó técnicas estadísticas para analizar el almacenamiento y la optimización de propiedades de las redes.

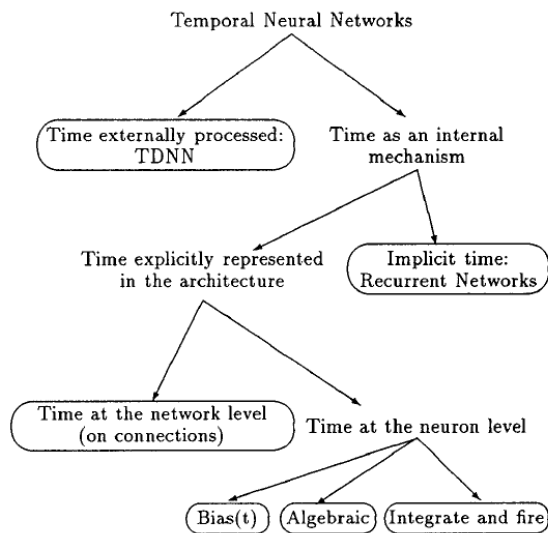


Figura 5: Redes neuronales temporales [4] (p.8)

Luego, los psicólogos Rumelhart e Hinton continuaron los estudios de modelos de redes neuronales relacionados con la memoria. También, aparecieron estudios acerca del procesamiento distribuido paralelo. Estos modelos son también llamados modelos conexionistas.

4. Arquitectura de las redes neuronales

Antes de entrar en detalle con ciertas disposiciones de la arquitectura de las redes neuronales, es necesario hacer notar la diferencia entre la inteligencia artificial conexionista y la simbólica.

Ambos enfoques buscan manejar aspectos complejos en búsquedas, representación de conocimiento y aprendizaje, en [7] se mencionan ciertas diferencias:

- Conexionista
 - o Búsqueda por relajación paralela.

- o Representación de conocimiento a través de conexiones y estructuras almacenadas como patrones distribuidos de activación.
- o Aprendizaje a través de retro-propagación, máquinas de Boltzman, entre otros.

- Simbólica

- o Búsqueda a través de estados de espacio transversal.
- o Representación de conocimiento a través de lógica de predicados, redes semántica, entre otros.
- o Aprendizaje a través de operadores macro, espacio de versiones, descubrimiento, entre otros.

En [3] se identificaron tres diferentes clases fundamentales de arquitectura de redes neuronales: las redes *feedforward* de capa simple, la redes *feedforward* multicapa y las redes recurrentes.

4.1. Redes *feedforward* de capa simple

En este clase las neuronas son organizadas en la forma de capas, en la forma simple se tiene una capa de entrada de nodos origen que proyecta dentro de la capa de salida de neuronas pero no viceversa, es decir, es una red no cíclica.

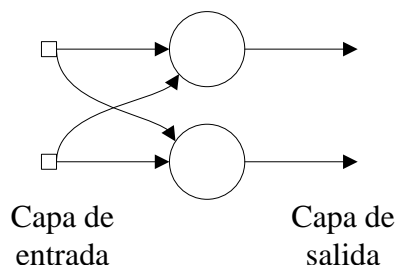


Figura 6: Redes neuronales de capa simple

4.2. Redes *feedforward* multicapa

Se diferencia de la de capa simple por la presencia de una o más capas ocultas, cuyos nodos computados son las correspondientemente llamadas neuronas ocultas o unidades ocultas. La función de las neuronas ocultas es intervenir entre la entrada externa y la red de salida de alguna manera útil.

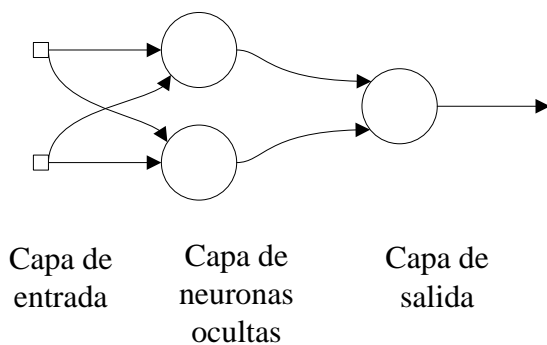


Figura 7: Redes neuronales multicapa

4.3. Redes recurrentes

Se diferencian a las anteriores clases porque tienen al menos un bucle o ciclo de realimentación, por ejemplo, una red recurrente es aquella que la salida de cierta neurona es la entrada de las demás neuronas.

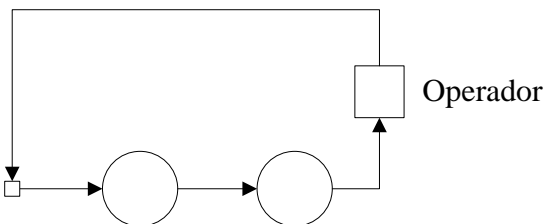


Figura 8: Redes recurrentes

5. Procedimientos y algoritmos

5.1. Funcionamiento

En [9], se hace referencia a unos elementos que computan funciones no lineales de suma de pesos ponderados en sus entradas, estos elementos son denominados “unidades lógicas de umbral” o TLU, y la red de TLU’s son conocidas como las redes neuronales, pero desde una perspectiva booleana funcional y de producción.

Se hace la diferencia entre dos tipos de redes, una llamada red *feedforward*, tal como fue vista en la arquitectura, que no tiene ciclos y las que si tiene son llamadas redes recurrentes.

A continuación se describen procedimientos para el entrenamiento multicapa, en redes *feedforward*.

5.2. Método de retro-propagación:

Dado un conjunto de entrada y salida en pares de vectores, este método computa un conjunto de pesos [7] para redes de tres niveles, que mapea entradas en sus correspondientes salidas.

En este método se computa el gradiente de un función cuadrática de error [9], en este caso el vector de peso sobre el gradiente es computado, debería ser un vector consistente de todos los pesos de la red. Es conveniente tomar las derivadas parciales de ϵ con respecto a varios pesos en grupos correspondientes al vector de pesos de los sigmoides individuales.

5.3. Aprendizaje de *perceptron* de incremento fijo

Dado un problema de clasificación con n entradas características y dos salidas clases, este algoritmo [7] computa un conjunto de pesos que causarán que un *perceptron* se almacene, siempre que la entrada falle en la primer salida clase. Es un algoritmo de búsqueda, iniciando con un estado inicial aleatorio y busca un estado de solución.

5.4. Aprendizaje competitivo

Dada una red compuesta de n valores binarios de unidades de entrada directamente son conectadas a cualquier número de unidades de salida [7], produce un conjunto de pesos, como las unidades de salida vienen activas de acuerdo a ciertas divisiones naturales de las entradas.

6. Aplicación de las redes neuronales

Existe un sinfín de aplicaciones relacionadas con las redes neuronales, en [7] se mencionan tres principalmente:

- Reconocimiento de voz: Este reconocimiento es una tarea de percepción difícil, el enfoque de redes neuronales presenta una red con palabras y sus pronunciaciones, para descubrir las regularidades y recordar las excepciones.
- Visión paralela: La visión humana es limitada, solo en el centro de la retina se mantiene una buena resolución, las redes sirven para construir modelos tridimensionales y proporcionar métodos para estudiar la visión paralela y adicionalmente la interpretación de la visión.

- Problemas combinatorios: También se pueden utilizar para resolver muchos problemas, por ejemplo, el problema del agente viajero, la coloración de grafos, etc.

En 1995 Schütze [10] comparó experimentalmente el indexado semántico latente (LSI) con una técnica basado en *chi* cuadrado con diferentes métodos de clasificación de aprendizaje como el análisis discriminante lineal, regresión logística y redes neuronales.

Orientando las redes neuronales como clasificadores. Los clasificadores de texto de redes neuronales son redes de unidades, donde las unidades de entradas representan términos, y las unidades de salida representan la categoría o categorías de interés, y el peso entre las unidades de enlaces de conexión representan las relaciones de dependencia.

Concluyendo que las redes neuronales y los clasificadores lineales en línea trabajan bastante bien, aunque son mejores trabajando conjuntamente con otros clasificadores basados en el análisis discriminante lineal y regresiones lineales.

Otra aplicación de las redes neuronales es la agrupación o *clustering*, entre algunas características [11] de las redes neuronales para agrupación de patrones están:

- Las redes neuronales artificiales procesan vectores numéricos y requieren de patrones para ser representados utilizando solamente características cuantitativas.
- Las redes neuronales artificiales son inherentes a arquitecturas de procesamiento distribuido y paralelo.

- Las redes neuronales pueden aprender por sus pesos adaptativos de sus interconexiones, es decir, pueden actuar como patrones normalizadores y selectores.

También en [2] se menciona otra aplicación, la cual, es el análisis de cambio de dimensiones de datos, especialmente en la fase de transformación en metales, siendo subjetivo cuando las medidas precisas son necesarias, para resolver este problema se desarrolla un método que es fundamentalmente justificado y fácilmente reproducible, llamado método de compensación.

Respecto a los problemas de control, en [12] se propone un método para evolucionar los conjuntos de redes neuronales, para realizar tareas de control en el establecimiento de refuerzo de aprendizaje. Este método es suficientemente general en principio que puede ser utilizado en conjunción con cualquier enfoque para la evaluación de una población de individuos que representan las políticas de un problema de control, por lo que es ampliamente utilizado.

7. Herramientas de software basadas en redes neuronales

7.1. Análisis y clasificación

EasyNN:

Es una herramienta para realizar un análisis de datos complejo, sirve para hacer predicciones, clasificaciones y proyecciones. Utiliza redes neuronales multicapa. Versión: 12.0d. Soporte para Windows. Licencia: Propietaria. Enlace web: <http://www.easynn.com/>

NeuroXL:

Tiene dos *add-in* para Microsoft Excel, uno es un clasificador de datos basado en redes neuronales y el otro es una herramienta de predicción también basada en redes neuronales para estimación de problemas. Con soporte para Windows. Versión XL. Licencia: Propietaria. Enlace web: <http://www.neuroxl.com/>

Tradecision:

Es una herramienta de análisis para toma de decisiones, orientada para usuarios especializados, contiene muchos indicadores e identifica patrones. Licencia: Propietaria. Enlace web: <http://www.tradecision.com/>

7.2. Ambientes de desarrollo

Neuro Laboratory:

Es un ambiente para el procesamiento de redes neuronales. Ayuda al procesamiento de señales, reconocimiento de imágenes, clasificaciones, predicciones y sistemas de decisión. Con soporte para Windows y Linux. Versión 1.0. Licencia: Propietaria. Enlace web: <http://www.scientific-soft.com/>

Synapse:

Es un ambiente de desarrollo para sistemas adaptativos, permite el análisis y procesamiento de datos y luego el diseño, entrenamiento, post-proceso y despliegue de sistemas adaptativos para resolución de tareas y problemas. Desarrolladores: Peltarion. Con soporte para Windows. Licencia EULA. Enlace web: <http://www.peltarion.com/products/synapse/>

7.3. Diseño y simulación

Neural Network Toolbox:

Es una extensión de MATLAB para diseñar, implementar y visualizar redes neuronales, provee soporte para varios paradigmas de red, simplificando la creación de redes y funciones. Con soporte para varias plataformas. Versión: 6.03. Licencia: Propietaria. Enlace web: <http://www.mathworks.com/products/neuralnet/>

Emergent:

Es un simulador de redes neuronales que permite la creación y en análisis de modelos complejos y sofisticados del cerebro en el mundo, con un énfasis en el análisis cualitativo y de enseñanza, también soporta un flujo de trabajo de investigadores de redes neuronales profesionales. Anteriormente llamado PDP++. Con soporte para Windows, OSX y Ubuntu. Desarrolladores: Universidad de Colorado. Versión 5.0.0. Licencia: GPL. Enlace web: http://grey.colorado.edu/emergent/index.php/Main_Page

Neural Lab:

Es un laboratorio avanzado para diseñar y entrenar redes neuronales artificiales en ingeniería, negocio, ciencias de la computación y tecnología. Con soporte para Windows. Desarrolladores: Universidad de Guanajuato. Versión 3.1. Licencia: Freeware. Enlace web: <http://www.dicis.ugto.mx/profesores/sledesma/documentos/index.htm>

Entre otras herramientas de software orientadas a la simulación y que tienen licencia Open Source están:

- Amygala
- Neuroph
- Simbrain

7.4. Machine Learning

Delve:

Es un ambiente estandarizado diseñado para evaluar el rendimiento de los métodos de aprendizaje basados en datos empíricos, realizando comparaciones con otros conjuntos de datos. Escrito en lenguaje C. Desarrolladores: Universidad de Toronto. Licencia: Libre, solamente si es para propósitos no comerciales. Enlace web: <http://www.cs.toronto.edu/~delve/>

Torch:

Es una biblioteca para *machine learning*, escrito en C++. Ensambla varios modelos, utiliza distribuciones, y tiene herramientas de reconocimiento de voz. Desarrolladores: Instituto de investigación IDIAP. Versión: 5. Licencia: BSD. Enlace web: <http://www.torch.ch/>

8. Conclusiones

Se realizó un pequeño estudio para describir de manera conceptual el área de las redes neuronales en la inteligencia artificial.

Las redes neuronales están compuestas por unidades que están interconectadas, estos enlaces tienen pesos asociados y existen funciones de activación para que la red alcance un estado estable.

La unidad básica de una red neuronal es una neurona, y sirve para el procesamiento de información en la operación de la red.

Las redes neuronales se pueden representar matemáticamente a través de grafos dirigidos.

Estas redes se basan en las redes de Hopfield, y en los *perceptrons*, según su arquitectura se clasifican en redes de capa simple, multicapa y redes recurrentes, la diferencia de estas últimas es la presencia de ciclos.

Existen muchas aplicaciones donde se utilizan las redes neuronales, entre las cuales se pueden mencionar el reconocimiento de voz, clasificación, agrupación, predicción, identificación de patrones y análisis de datos.

Se presentó una breve reseña de algunas herramientas de software disponibles en la web que utilizan redes neuronales.

9. Referencias bibliográficas

- [1] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 2nd ed. United States of America: Pearson Education, Inc., 2003.
- [2] H. Bhadeshia, "Neural Networks and Information in Materials Science," University of Cambridge, 2008.
- [3] S. Haykin, *Neural Networks. A Comprehensive Foundation*, 2nd ed. United States of America: Prentice Hall, Inc., 1999.
- [4] J. C. Chappelier and A. Grumbach, "Time in Neural Networks," *SIGART Bulletin*, vol. 5, no. 3, pp. 3-11, Jul. 1994.
- [5] P. J. S. Jr. and A. S. Elmaghraby, "Selection of a Neural Networks System for Visual Inspection," University of Louisville, 1989.
- [6] J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," in *Proc. NatL Acad. Sci. USA, Biophysics*, United States of America, 1982, pp. 2554-2558.
- [7] E. Rich and K. Knight, *Artificial Intelligence*, 2nd ed. United States of America: McGraw-Hill, Inc., 1991.
- [8] F. Ronsenblatt, "The perceptron: A Probabilistic Model for Information Storage and Organization in the Brain.," *Psychological Review*, vol. 656, pp. 386-408, 1958.
- [9] N. Nilsson, *Artificial Intelligence: A New Synthesis*. United States of America: Morgan Kaufmaann Publisher, Inc., 1998.
- [10] F. Sebastiani, "Machine Learning in Automated Text Categorization," in *ACM Computing Surveys, Vol. 34, No. 1*, 2002, pp. 1-47.
- [11] A. Jain, M. Murty, and P. Flynn, "Data Clustering: A Review," *ACM Computing Surveys*, vol. 31, no. 3, pp. 264-323, Sep. 1999.
- [12] D. Pardoe, M. Ryoo, and R. Miikkulainen, "Evolving Neural Network Ensembles for Control Problems," in *GECCO'05*, Washington, DC, USA., 2005, pp. 1379-1384.