INTELIGENCIA ARTIFICIAL

JAIRO SALAZAR

CRISTIAN RESTREPO

PROFESOR:

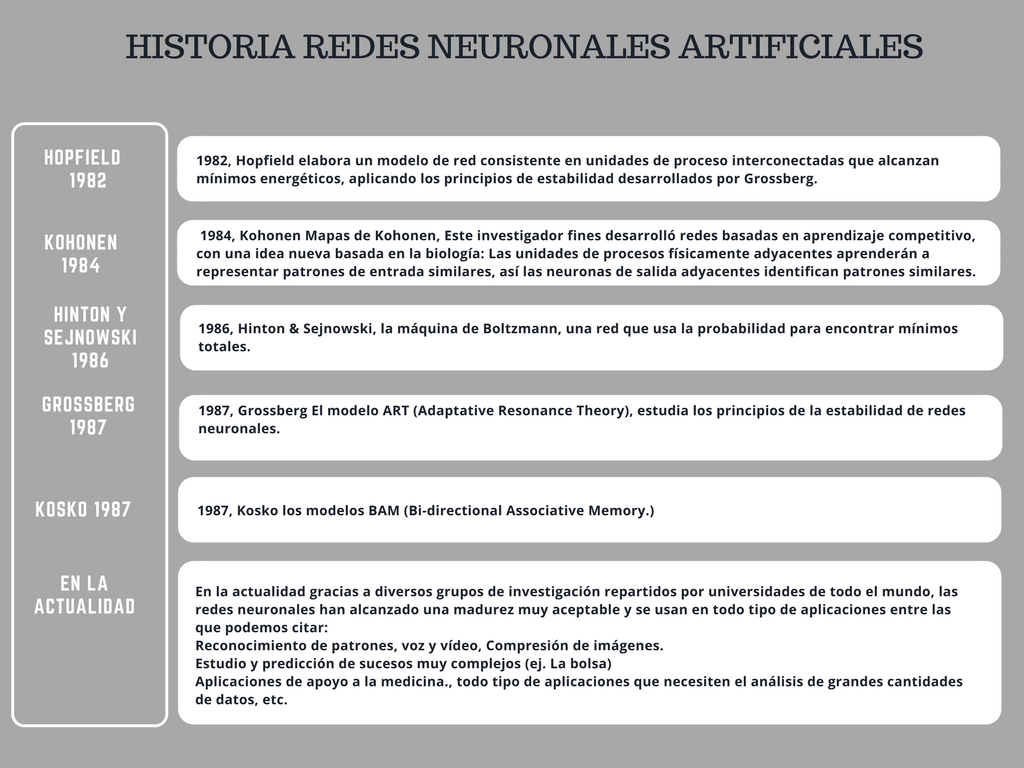
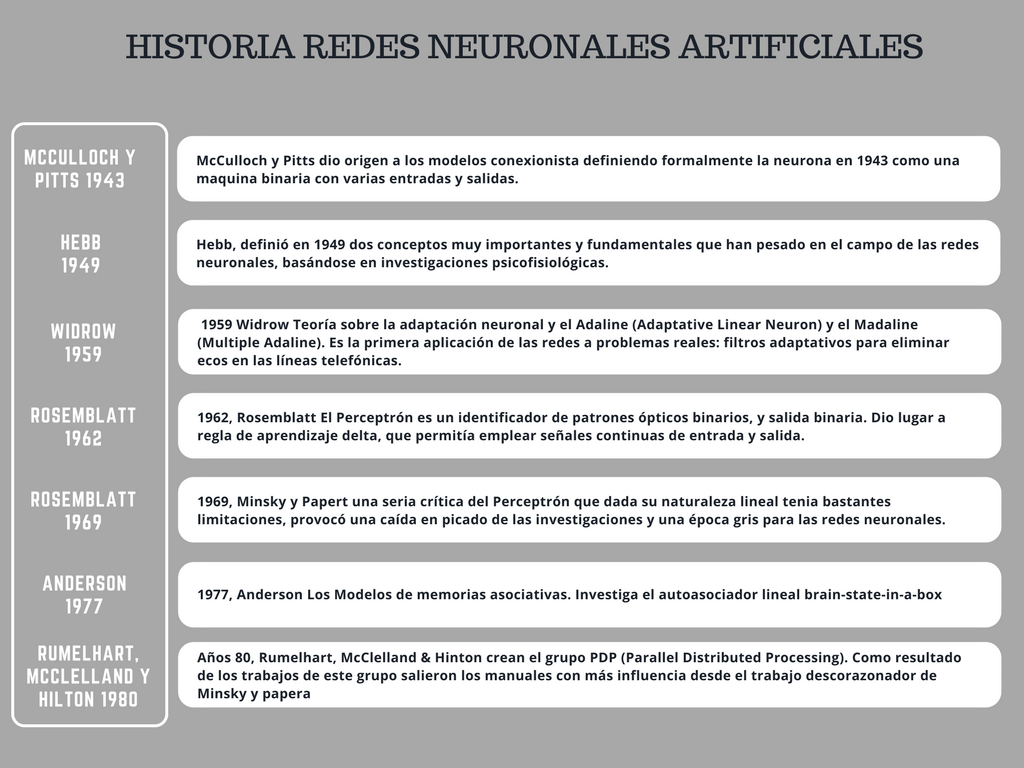
CARLOS LONDOÑO

CORPORACION DE ESTUDIOS TECOLOGICOS DEL NORTE DEL VALLE

TECNOLOGIA EN SISTEMA DE INFORMACION

CARTAGO 28 FEBRERO 2018

Taller N 3

1. Consultar la historia de las redes neuronales artificiales, crear un mapa conceptual, que permita evidenciar los casos más importantes.
2. Nombra 6 o más ventajas y desventajas que tiene el uso de las redes neuronales artificiales para la solución de problemas.

**Ventajas:**

* El aprendizaje de las redes neuronales artificiales tiene la habilidad de aprender mediante una etapa que se llama etapa de aprendizaje. Esta consiste en proporcionar a las redes neuronales artificiales datos como entrada a su vez que se le indica cuál es la salida esperada.
* La auto organización de una red neuronal artificial crea su propia representación de la información en su interior.
* La tolerancia a fallos se da debido a que una red neuronal artificial almacena la información de forma redundante, ésta puede seguir respondiendo de manera aceptable aun si se daña parcialmente.
* La flexibilidad en una red neuronal artificial puede manejar cambios no importantes en la información de entrada, como señales con ruido u otros cambios en la entrada.
* El tiempo real en la estructura de una red neuronal artificial es paralela, por lo cual, si esto es implementado con computadoras o en dispositivos electrónicos especiales, se pueden obtener respuestas en tiempo real.
* Las redes neuronales artificiales son excelentes como reconocedores de patrones y pueden ser usadas donde las técnicas tradicionales no funcionan. Las redes neuronales pueden manejar excepciones y entradas de datos anormales, muy importante para sistemas que manejan un amplio rango de datos (sistemas de radar y sonar, por ejemplo). Muchas redes neuronales son biológicamente plausibles, lo que significa que pueden proveer pistas de cómo trabaja el cerebro según progresen.

**Desventajas:**

* Complejidad de aprendizaje para grandes tareas, cuantas más cosas se necesiten que aprenda una red, más complicado será enseñarle.
* Tiempo de aprendizaje elevado. Esto depende de dos factores: primero si se incrementa la cantidad de patrones a identificar o clasificar y segundo si se requiere mayor flexibilidad o capacidad de adaptación de la red neuronal para reconocer patrones que sean sumamente parecidos, se deberá invertir más tiempo en lograr que la red converja a valores de pesos que representen lo que se quiera enseñar.
* No permite interpretar lo que se ha aprendido, la red por si sola proporciona una salida, un número, que no puede ser interpretado por ella misma, sino que se requiere de la intervención del programador y de la aplicación en si para encontrarle un significado a la salida proporcionada.
* Elevada cantidad de datos para el entrenamiento, cuanto más flexible se requiera que sea la red neuronal, más información tendrá que enseñarle para que realice de forma adecuada la identificación.

1. Nombre 10 aplicaciones de las redes neuronales.

* En la construcción de modelos explicativos, que ayudan a explorar conjuntos de datos en busca de variables o grupos de variables pertinentes.
* En su capacidad para aproximar cualquier función continua, haciendo posible que el analista no necesite tener ninguna hipótesis sobre el modelo subyacente.
* Son capaces de determinar relaciones no lineales entre un conjunto de datos, asociando patrones de entrada o salidas correspondientes.
* Los tipos de aprendizaje disponibles pueden utilizarse para tareas de predicción y clasificación.
* Los modelos supervisados y no supervisados pueden ser aplicados para extraer y cancelar ruido de las señales.
* Una vez que la red ha sido entrenada y probada puede adaptarse por sí misma a los cambios.
* Una aproximación basada en redes neuronales artificiales puede aprender los modelos específicos de cada sistema de red y proporcionar aproximaciones aceptables de los sistemas.
* Explotación de bases de datos.
* Reconocimiento de caracteres escritos.
* Modelado de sistemas para automatización y control.
* Creación de armas inteligentes.

1. ¿Qué son funciones de activación, cuales existen y para cuales redes neuronales se aplican?

Las funciones de activación son una idea tomada de la Biología y que trata de generalizar el hecho de que las neuronas biológicas no son meros sumadores y transmisores de impulsos, sino que tienen una suerte de mecanismo que decide si se activan o no, en función de la entrada que reciben. Es al activarse cuando envían una señal a través de su axón a otras neuronas que la siguen en la cadena de procesamiento.

Desde el punto de vista de las redes neuronales artificiales, las funciones de activación cumplen dos papeles muy importantes y son una característica con margen para hacer diseños novedosos de redes neuronales artificiales.

* Función de umbral
* Función sigmoidal o función logística t(x) = 1 / (1 + *e-cx*)
* Tangente hiperbólica t(x) = (*e-cx* - *e-cx*) / (*e-cx* + *e-cx*)
* Función gausiana.

1. Perceptrón:
2. Historia

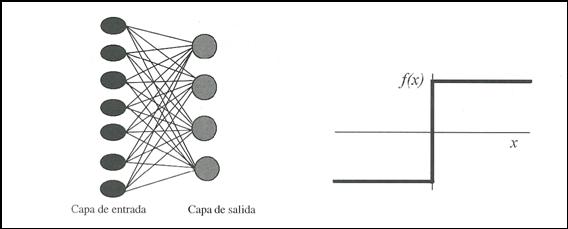
En 1969, Minsky y Papert, demuestran que el perceptrón simple y ADALINE no puede resolver problemas no lineales (por ejemplo, XOR). La combinación de varios perceptrones simples podría resolver ciertos problemas no lineales, pero no existía un mecanismo automático para adaptar los pesos de la capa oculta. Rumelhart y otros autores, en 1986, presentan la "Regla Delta Generalizada" para adaptar los pesos propagando los errores hacia atrás, es decir, propagar los errores hacia las capas ocultas inferiores. De esta forma se consigue trabajar con múltiples capas y con funciones de activación no lineales. Se demuestra que el perceptrón multicapa es un aproximador universal. Un perceptrón multicapa puede aproximar relaciones no lineales entre los datos de entrada y salida. Esta red se ha convertido en una de las arquitecturas más utilizadas en el momento.

1. Fórmula matemática, explique sus términos

El perceptrón simple es un modelo neuronal unidireccional, compuesto por dos capas de neuronas, una de entrada y otra de salida. La operación de una red de este tipo, con n neuronas de entrada y m neuronas de salida, se puede expresar de la siguiente forma:

http://grupo.us.es/gtocoma/pid/pid10/RedesNeuronales_archivos/image138.gif

1. Dibuje la estructura

  
*Perceptrón simple y función de transferencia de su neurona.*

1. ¿Para qué se usa?

El perceptrón simple sólo sirve para clasificar problemas linealmente separables, cosa que ya se podía hacer mediante métodos estadísticos, y de una forma mucho más eficiente.

Concluimos añadiendo que el perceptrón simple está formado por dispositivos de umbral y, por tanto, son útiles para la representación de funciones booleanas.

1. ¿Cuál es su función de activación?
2. ¿Cómo se entrena un perceptrón, indique el nombre del algoritmo y sus pasos?

* El algoritmo de aprendizaje es de tipo supervisado.
* En el proceso de entrenamiento, el Perceptrón se expone a un conjunto de patrones de entrada, y los pesos de la red son ajustados de forma que al final del entrenamiento se obtenga las salidas esperadas para cada uno de esos patrones de entrada.
* A continuación, el algoritmo de ajuste de pesos para realizar el aprendizaje de un Perceptrón (aprendizaje por corrección de error).
* Inicialización de los pesos y del umbral

Inicialmente se asignan valores aleatorios a cada uno de los pesos wi de las conexiones y al umbral 

* Presentación de un nuevo par (Entrada, salida esperada)

Patrón de entrada Xp = ( x1, x2, x3, …, xn), salida esperada d (t).

Cálculo de salida actual



siendo f la función de transferencia escalón.

* Adaptación de los pesos



a es un factor de ganancia en el rango 0 a 1.

1. Nombre 5 ejemplos donde se evidencie el uso del perceptrón.
2. Adaline:
3. Historia

La Red Adaline fue desarrollada en el 1960 por Bernard Widrow y su estudiante Marcian Hoff de la universidad de Stanford.

ADALINE proviene de Adaptive Lineal Element (Elemento Lineal Adaptativo), pero antes de que se le diera este nombre esta red sufrió un cambio ya que primeramente se llamaba Adaptive Lineal Neuron (Neurona Linear Adaptiva), dicho cambio se dio por que la Red Adaline es un dispositivo que consta de un único elemento de procesamiento, como tal no es técnicamente considerada una red neuronal.

Adaline fue desarrollada para el reconocimiento de patrones binarios, por ejemplo, predecir el siguiente bit en una línea telefónica.

1. Fórmula matemática, explique sus términos

Generalmente se compone de una sola capa de n neuronas (por tanto n valores de salida) con m entradas con las siguientes características:

Las m entradas representan un vector x de entrada que pertenece al espacio R^m.

Por cada neurona, existe un vector w de pesos sinápticos que indican la fuerza de conexión entre los valores de entrada y la neurona. En la práctica representan la ponderación de cada entrada sobre la neurona.

Una constante 0(theta).

La salida y de la neurona se representa por la función de activación, que se define como 

1. Dibuje la estructura
2. ¿Para qué se usa?

La principal aplicación de las redes de tipo Adaline se encuentra en el campo de procesamiento de señales. Específicamente en el diseño de filtros, estas son capaces de eliminar ruido en señales portadoras de información.

Otra aplicación que tiene esta red es la de los filtros adaptivos.

Un Adaline tiene la suficiente capacidad de predecir el valor de una señal en el instante (t+1) si se conoce el valor de la misma en los p instantes anteriores (p es >0 y su valor depende del problema).

Predecir el valor futuro de una señal a partir de su valor actual.

El error predicción va a variar, es decir la predicción va a ser mayor o menor dependiendo de la señal que se va a predecir. Si dicha señal corresponde a una serie de temporal el Adaline, transcurrido un tiempo, estará apta para dar predicciones correctas.

1. ¿Cuál es su función de activación?

Las Redes ADALINE son redes muy similares al Perceptrón con la diferencia de que su función de activación es lineal en lugar de ser un limitador fuerte como es el caso del Perceptrón, estas presentan la misma limitación del Perceptrón respecto al tipo de problemas que pueden resolver, ya que ambas redes solo pueden resolver problemas linealmente separables.

Son redes de aprendizaje supervisado que usan la regla de Widrow – Hoff para dicho aprendizaje o también denominada regla Delta. El algoritmo que estas usan es el LMS (Least Mean Square) siendo este más eficiente que la regla de aprendizaje del Perceptrón puesto que minimiza el error medio cuadrático.

Adaline está limitada a una única neurona de salida, un vector x como su entrada y un número real y como su salida.

1. ¿Cómo se entrena una red adaline, indique el nombre del algoritmo y sus pasos?

La salida lineal obtenida del ALC es aplicada a un Conmutador Bipolar.

El Umbral se representa a través de una conexión ficticia de peso W0 (b).

La red Adaline puede tomar valores continuos.

1.- Inicializar los pesos en forma aleatoria

2.- Introducir un patrón de entrada

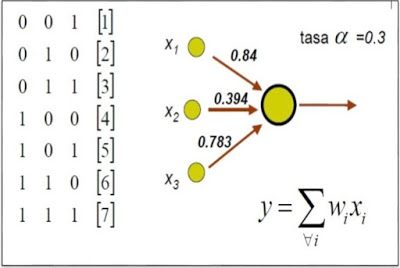
3.- Calcular la salida (y), compararla con la deseada (d) y obtener la diferencia (dp - yp)

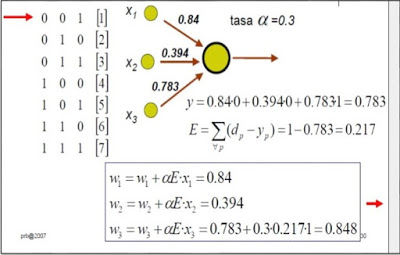
4.- Multiplicar el resultado del paso anterior por la entrada correspondiente a cada uno de los pesos y ponderarla por la tasa de aprendizaje.

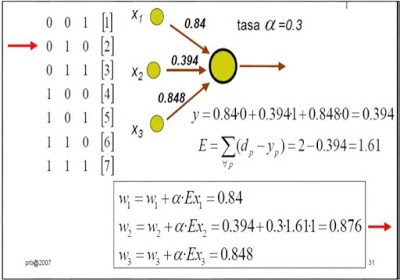
5.- Actualizar los pesos, sumando al valor antiguo la cantidad obtenida en el paso anterior

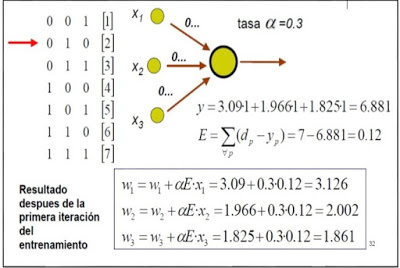
6.- Si no se ha cumplido el criterio de parada, regresar al paso 2, si se ha acabado todos los patrones repetir el algoritmo.

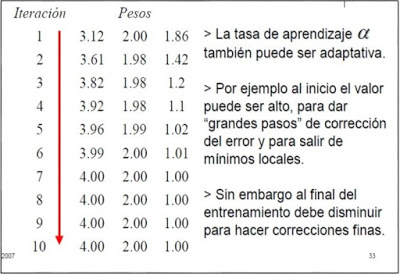
1. Nombre 5 ejemplos donde se evidencie el uso del perceptrón.

**[](http://3.bp.blogspot.com/-THsUIgnjomY/VZ3WvLgkS-I/AAAAAAAAAJg/fl9uQ6TOpvo/s1600/Diapositiva2.JPG)**

**[](http://2.bp.blogspot.com/-DYQPeMOuFG8/VZ3W2ZtpT9I/AAAAAAAAAJo/S0IlesS3LcA/s1600/Diapositiva3.JPG)**

**[](http://2.bp.blogspot.com/-iZvJ9tNByzA/VZ3W9cP9y-I/AAAAAAAAAJw/dX_hvq7GMNU/s1600/Diapositiva4.JPG)**

**[](http://4.bp.blogspot.com/-GXNq0hamYyo/VZ3XJbS-9CI/AAAAAAAAAJ4/htekvHglWPs/s1600/Diapositiva5.JPG)**

**[](http://3.bp.blogspot.com/-8rU4I5GQfRU/VZ3XPcxckkI/AAAAAAAAAKA/3rOmZUG20pU/s1600/Diapositiva6.JPG)**

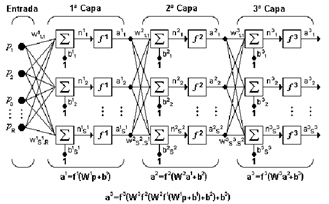
1. ¿Qué es y para que se usan las redes de retroprogamacion (backpropagation)?

La propagación hacia atrás de errores o retropropagación es un método de cálculo del gradiente utilizado en algoritmos de aprendizaje supervisado utilizados para entrenar redes neuronales artificiales. El método emplea un ciclo propagación – adaptación de dos fases. Una vez que se ha aplicado un patrón a la entrada de la red como estímulo, este se propaga desde la primera capa a través de las capas siguientes de la red, hasta generar una salida. La señal de salida se compara con la salida deseada y se calcula una señal de error para cada una de las salidas.

Las salidas de error se propagan hacia atrás, partiendo de la capa de salida, hacia todas las neuronas de la capa oculta que contribuyen directamente a la salida. Sin embargo, las neuronas de la capa oculta solo reciben una fracción de la señal total del error, basándose aproximadamente en la contribución relativa que haya aportado cada neurona a la salida original. Este proceso se repite, capa por capa, hasta que todas las neuronas de la red hayan recibido una señal de error que describa su contribución relativa al error total.

La importancia de este proceso consiste en que, a medida que se entrena la red, las neuronas de las capas intermedias se organizan a sí mismas de tal modo que las distintas neuronas aprenden a reconocer distintas características del espacio total de entrada. Después del entrenamiento, cuando se les presente un patrón arbitrario de entrada que contenga ruido o que esté incompleto, las neuronas de la capa oculta de la red responderán con una salida activa si la nueva entrada contiene un patrón que se asemeje a aquella característica que las neuronas individuales hayan aprendido a reconocer durante su entrenamiento.

1. Explique el algoritmo backprogramation, usando imágenes.

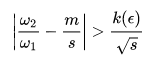


Puede notarse que esta red de tres capas equivale a tener tres redes tipo Perceptrón en cascada; la salida de la primera red, es la entrada a la segunda y la salida de la segunda red es la entrada a la tercera. Cada capa puede tener diferente número de neuronas, e incluso distinta función de transferencia.

1. Haciendo uso del lenguaje de programación Python graficar las siguientes funciones:
2. Función escalón
3. Función lineal y mixta
4. Función tangente hiperbólica
5. Función sigmoidal
6. Función gauss
7. ¿Qué es el teorema de kolmogorov, explicar?

El teorema de Kolmogórov–Arnold–Moser o teorema KAM es un resultado de sistemas dinámicos sobre la persistencia de movimientos cuasiperiódicos. Este teorema resuelve parcialmente el problema de los divisores pequeños (que origina problemas de convergencia en sistemas con múltiples frecuencias). El teorema explica cómo se modifica el aspecto de las trayectorias de un sistema integrable bajo pequeñas perturbaciones.

El teorema KAM establece que, si un sistema está sometido a una pequeña perturbación no lineal, algunos toros serán deformados y otros destruidos. Los que sobreviven son aquellos que tienen un cociente de frecuencias suficientemente irracional. Es decir, se destruyen aquellos cuyo cociente de frecuencias se acerca más a un número racional, dados por la relación.

v

Con  El último toro en destruirse es el más irracional de todos (el que guarda mayor semejanza con el número áureo). Informalmente el teorema establece que:

"Para perturbaciones suficientemente pequeñas, casi todos los toros invariantes se preservan [en el sistema perturbado] (excluyendo aquellos con vectores de frecuencia racionales)"