

NUEVOS PARADIGMAS REDES NEURONALES Y EDUCACIÓN

.-Redes Neuronales Artificiales vs Redes Neuronales Biológicas

- 1) Comparación entre humanos y máquinas
- 2) Las neuronas y el aprendizaje
- 3) Activación de las neuronas y su representación en las redes neuronales

II.- Redes neuronales

- 1) Definiciones
- 2) Introducción de las redes neuronales
- 3) Back-Propagation
- 4) Problemas relacionados con el modelo neuronal

III) Redes neuronales multi-nivel de alimentación hacia adelante

- 1) Funcionamiento
- 2) Interpretación de entradas y salidas
- 3) Entrenamiento
- 4) Características importantes de muestras para entrenamiento
- 5) Problemas
- 6) Entrenamiento no supervisado
- 7) Modelos de entrenamiento y topologías de red

IV) Aplicaciones de redes neuronales

- 1) Control de la eficiencia de una máquina
- 2) Reconocimiento de firmas
- 3) Reconocimiento de blancos mediante sonares
- 4) Predicción en la bolsa de valores
- 5) Otras

V) Ejemplo del entrenamiento de una red neuronal

I.- Redes Neuronales Artificiales vs Redes Neuronales Biológicas

1) *Comparación entre humanos y máquinas*

El aparato computacional mas poderoso conocido por el hombre es el cerebro humano. Un niño de tres años puede realizar fácilmente labores que superan por mucho las capacidades de las computadoras mas sofisticadas: reconociendo docenas de caras y cientos de objetos desde diferentes ángulos, en condiciones de luz distintas; manipulando un ambiente complejo; entendiendo y usando un vocabulario complejo de lenguaje y gestos. Han sido gastados enormes esfuerzos en desarrollo para intentar reproducir versiones limitadas de algunas de estas capacidades en las computadoras, con pobres resultados.

Una computadora puede realizar en segundos cálculos aritméticos que a un ser humano le llevarían años. La aritmética es difícil para los humanos e imposible para los animales, dicha habilidad se considera tradicionalmente como un signo de gran inteligencia. Entonces, ¿que causa la disparidad entre las habilidades del hombre y la maquina?

La diferencia obvia yace en la arquitectura fundamental de la computadora y del cerebro. Computadoras tradicionales están basadas en la arquitectura de Von Neumann: un simple procesador que puede realizar operaciones simples aritméticas, lógicas y condicionales una a la vez, y una gran memoria, a tazas de millones por segundo o mas.

El cerebro humano, en contraste, consiste de un gran numero de células especializadas llamadas neuronas, interconectadas masivamente (se estima que hay en el orden de diez billones de neuronas en el cerebro humano con un promedio de miles de conexiones por cada una). Estas neuronas son lentas (realizando cientos en lugar de millones de operaciones por segundo).

Las Redes Neuronales Artificiales (ANN, por sus siglas en ingles o simplemente Redes Neuronales) son modelos computarizados inspirados en la estructura a bajo nivel del cerebro. Consisten en grandes cantidades de unidades de procesamiento sencillas llamadas neuronas, conectadas por enlaces de varias fuerzas. Redes Neuronales también pueden ser construidas con hardware especial o simuladas en computadoras normales. Sin embargo, el hardware neuronal especializado no es común, así que la simulación es la norma.

2) *Las neuronas y el aprendizaje*

El cerebro contiene un gran numero de células especializadas llamadas neuronas. Una neurona tiene tres partes: un cuerpo celular, una fina estructura de entrada (Dendrita) y una gran estructura de salida (Axon). Un buen mnemónico es que dendrita=Input (Entrada) y axon=output (salida). Las terminales de los axones terminan en synapsis:

estas son conexiones que casi todas los dendritas de otras neuronas. Señales electroquímicos se propagan por neuronas desde sus synapsis hacia otras neuronas.

Esta conducta simple es modificada en varias formas:

a) Una neurona solo produce una señal (se dispara) si la señal de entrada supera una determinada cantidad en un periodo corto de tiempo.

b) Las synapsis varían en fuerza; algunas son buenos conectores permitiendo una señal fuerte, y otras una señal débil.

c) Las synapsis pueden ser: excitatorias o inhibitorias. Una synapsis excitatoria añade una señal a la dendrita. En contraste, la inhibitoria reduce la señal de la dendrita.

Ahora se cree que gran parte de la actividad cerebral es controlado por conexiones introductoras/removedoras entre neuronas, y alterando las fuerzas sinápticas de las conexiones. Por ejemplo, asumamos que dos neuronas representan dos conceptos: Comida y Campana. La neurona Comida transmite cada vez que la comida esta disponible y la neurona Campana cuando suena un campana. Si hay cena la campana suena, de ahí que hay una conexión muy importante entre estos dos conceptos. El aprendizaje de Hebb postula que la fuerza synaptica entre neuronas se incrementa si representan conceptos asociados. Cada vez que la campana suena y la comida aparece, la conexión entre estas dos neuronas aumenta. Consecuentemente, el cerebro aprende a asociar la campana con comida. Ahora si hay un numero de conceptos que indican la proximidad de comida (hambre, olor) que pueden combinarse y "pesarse" de acuerdo a su importancia relativa para determinar si la comida puede aparecer. La neurona de comida quizá pueda activarse solos si coinciden un número de conceptos relacionados, por ejemplo, si la campana suena cuando no hay comida, puede no activarse.

3) Activación de las neuronas y su representación en las redes neuronales

Matemáticamente podemos representar una neurona simplificada por un valor (que debe ser superado para que se active) y una lista de sus sinapsis y sus fuerzas asociadas. Las señales de entrada a una neurona son multiplicadas por sus fuerzas("pesos") asociadas y después se suman. El resultado se llama el nivel de activación de la neurona. Si el nivel de activación supera el valor de la neurona, esta se activa, y una señal se envía a cada neurona que tiene conectada.

Se ha postulado que paquetes de neuronas deben compartir esencialmente entradas de otros paquetes, así que la conducta de neuronas individuales es irrelevante. Los paquetes de neuronas mas que neuronas individuales necesitan ser modelados. En una red neuronal, se puede tomar que una neurona sencilla representa un paquete de neuronas reales para enfatizar que el modelo de neuronas usado hasta ahora es extremadamente simple con respecto al del cerebro humano.

II.- Redes nueronales

1) Definiciones

Estudio de Redes Neuronales de DARPA (1988, AFCEA International Press, p. 60):

"Una red neuronal es un sistema compuesto de muchos elementos procesadores simples operando en paralelo, cuya función es determinada por la estructura de la red, fuerza en las conexiones y el procesamiento realizado por los elementos computacionales en los nodos."

Haykin, S. (1994, Neural Networks: A comprehensive foundation, NY, McMillan, P.2):

"Una red neuronal es un procesamiento distribuido masivamente paralelo que tiene una tendencia natural para almacenar conocimiento empírico y hacerlo disponible para el uso. Recuerda al cerebro en dos aspectos:

- 1.- Conocimiento se adquiere por la red a través de un proceso de aprendizaje.
- 2.- Las conexiones interneurónicas se conocen como pesos synapticos y se usan para almacenar el conocimiento."

Zurada, J.M. (1992, Introduction to artificial neural systems, Boston: PWS Publishing Company, P. XV):

"Sistemas de redes neuronales artificiales, o redes neuronales son sistemas celulares físicos que puedan adquirir, almacenar y usar conocimiento empírico."

Las redes neuronales son una forma de un sistema computarizado multi-proceso con:

- Elementos de procesamiento sencillos.
- Alto grado de interconexión.
- Mensajes simples escalares.
- Interacción adaptable entre elementos.

2) Introducción de las redes neuronales

En los 50's y 60's, se hicieron varios intentos para adaptar los modelos de Redes Neuronales para que se diera el aprendizaje. Rosenblatt diseñó el Perceptron. Este contenía tres tipos de neuronas: sensoriales, asociativas y de respuesta. Las sensoriales tomaban entradas de fuera de la red, las unidades de respuesta propagaban señales afuera de la red al mundo externo, y las asociativas eran meramente internas. La distinción entre estas tres neuronas es importante aún sin embargo ahora se refieren como unidades de entrada, de salida y ocultas. Rosenblatt desarrolló métodos para alterar los niveles synapticos de forma que la red aprendiera a reconocer clases de entradas. Por ejemplo, produjo una red que aprendió a responder a líneas verticales, pero no a horizontales (se sabe que neuronas especializadas en la visión actúan de esta forma). Como muchas Redes Neuronales posteriores, el rasgo más importante del Perceptron de Rosenblatt que aprendió a clasificar sus entradas "por ejemplo"; lo cual contrasta con la clásica ciencia computacional donde el programador escribe un programa que le dice a la computadora como clasificar sus entradas.

En 1969, Minsky y Papert publicaron un libro que influyó mucho el pensamiento sobre redes neuronales. En primer lugar demostraron que para hacer algunas tareas importantes de clasificación en geometría requerían un incremento arbitrario en el tamaño del Perceptron mientras se iba incrementando el tamaño de la retina. En segundo lugar

demonstraron que los Perceptrons eran incapaces de aprender a resolver cualquier problema linealmente inseparable, puesto que problemas muy simples son linealmente inseparables (por ejemplo el XOR) este fue un grave problema.

Después de los 80's el campo de inteligencia artificial se hizo muy popular. En Inteligencia Artificial, la inteligencia se modela top-down, con algoritmos diseñados para modelar procesos mentales de alto nivel como la asociación de conceptos, deducción, inducción y razonamiento.

3) Back-Propagation

Las Redes Neuronales se pusieron de moda en 1986 cuando Rumelhart y McClelland. Ellos mostraron que algunas de las distinciones que son imposibles para Perceptrons simples pueden ser resueltas por redes multi-nivel con funciones de activación no-lineales, usando un procedimiento simple de entrenamiento: los algoritmos back-propagation. Back-propagation y sus derivados son los métodos mas importantes y difundidos de entrenamiento de redes usados en la actualidad.

El término de back-propagation causa mucha confusión, estrictamente hablando, back-propagation se refiere al método para computar el gradiente de error de una red feed-forward, una aplicación de vanguardia, pero elegante de la regla de la cadena de calculo elemental. Por extensión, back-propagation se refiere al método de entrenamiento que usa back-propagation para calcular el gradiente.

Básicamente el entrenamiento de estas redes consiste de :

- 1) La pasada hacia adelante (forward pass): las salidas son calculadas y el error en las unidades de salida es calculado.
- 2) La pasada hacia atrás (backward pass): el error de las unidades de salida es usado para alterar los pesos en las unidades de salida. Luego el error en los nodos ocultos es calculado (mediante la propagación hacia atrás (back-propagation) del error en las unidades de salida mediante los pesos), y los pesos en los nodos ocultos son alterados usando estos valores.

Para cada dato que se quiera aprender se ejecutan una pasada hacia adelante y otra pasada hacia atrás. Estos se repite una y otra vez hasta que el error este a una nivel suficientemente bajo.

La arquitectura típica de una red neuronal de back-propagation es:

4) Problemas relacionados con el modelo neuronal

Primero, si las señales llegan a la neurona a tiempos distintos, se deben sumarlas, y si es así como se deben sumar?. Si las señales viajan a diferentes tasas entre diferentes conexiones, habrá diferencias en los tiempos seguramente. Deben las neuronas esperar a que lleguen un cierto número de señales para ejecutar la función de activación?, si es así cuanto deben de esperar?. Las redes biológicas deben de tratar con estos problemas así que las redes artificiales deben de tomarlos en cuenta. Sin embargo, los modelos artificiales con los que se trataran hacen a un lado este problema insistiendo en que la red se actualiza en una secuencia de pasos discretos: en cada paso, se propagan las señales,

se calculan los niveles de activación y se determinan las salidas, a través de la red entera de una manera simultánea (o cuando menos en un orden bien definido).

Segundo. que neurona debe estar conectada con cual? En el caso mas general cada neurona puede tener conexiones con cualquier otra, lo que incluye conexiones de la neurona consigo misma. Un sistema donde la salida de un componente puede influenciar su entrada se dice que exhibe feed-back (retroalimentación); los sistemas con feed-back son notablemente difíciles de analizar y predecir. Una Red Neuronal con conexiones feed-back recibe el nombre de Red Recurrente. Las Redes Recurrentes tienen propiedades no deseadas; por ejemplo, las señales se pueden propagar por la red, indefinidamente, siempre cambiando. En este caso, es difícil ver como se puede interpretar que la red ha decidido algo, ya que la decisión requiere cierta estabilidad. Se ha hecho un gran esfuerzo en investigación en redes neuronales para conseguir estructuras de red que consigan estabilidad. La estabilidad puede introducirse si se deshace cualquier conexión feed-back o limitando el feed-back para hacer predecible y estable el comportamiento de la red.

III) Redes neuronales multi-nivel de alimentación hacia adelante

(Multi-Layer feedforward Neural Networks)

1) 1) *Funcionamiento*

El modelo de red multi-nivel feedforward (el concepto de feedforward quiere decir libre de ciclos, generalmente las conexiones permitidas son entre las capas de entrada y la primera capa oculta, de la primera oculta a la segunda, y de la ultima capa oculta a la de salida) es el mas popular, el cual se arregla en un numero de capas secuenciales; cada neurona en cada capa se conecta a cada neurona del siguiente nivel. La primera capa es llamada capa de entrada y recibe señales del mundo externo. La ultima es llamada capa de salida y propaga señales (decisiones de la red) al mundo exterior. Las otras capas son llamadas ocultas porque no se acceden directamente desde el mundo exterior ya que están envueltas exclusivamente en decisiones. El comportamiento de la red se basa en lo siguiente: Primero, las activaciones de las unidades de entrada son valuadas de acuerdo al problema que se quiere resolver. Después, las señales se propagan de la capa de entrada a la siguiente, con cada unidad realizando una suma ponderada de sus entradas y aplicando su función de activación. El proceso se repite hasta que la capa de salida se alcance. Las activaciones de la salida representa la respuesta de la red a las entradas que recibe.

2) 2) *Interpretación de entradas y salidas*

El significado preciso de los valores de entrada y salida depende de la aplicación por la que fue diseñada la red. Por ejemplo, si la red se pretende que prediga valores de mercado, las entradas pueden ser varias medidas financieras (precios anteriores, inflación, moneda) y las salidas pueden representar la predicción de la red respecto a los precios de mañana. Si la red predice una falla en un aeroplano, las entradas pudieron ser temperatura, nivel de aceite, altitud y la edad del avión,



con la salida siendo un valor entre cero y uno (0=tranquilo a 1=panico). En general, cualquier numero de entradas y cualquier numero de salidas puede ser especificado.

3)Entrenamiento

La parte mas importante de la modulación de la red neuronal es determinar formas para ajustar los pesos de forma que la red realice un buen mapeo entre entradas y salidas: a esto se le llama "entrenamiento" de la red.

El mapeo de redes neuronales es tan complejo que aun cuando una red ha sido bien entrenada para resolver un problema, el usuario puede ser incapaz de entender como la red logra esto. Lo que se requiere son métodos de entrenamiento automático para el mapeo de entrada-salida. Un numero de muestras son presentadas a la red y un procedimiento de entrenamiento es usado para ajustar los pesos para que aprenda el mapeo. Las muestras de entrenamiento consisten de un conjunto de entradas juntas con la salida esperada.

Entrenar una red neuronal es, en muchos de los casos, un ejercicio de optimización numérica de una función usualmente no-lineal. Se usa el entrenamiento de redes neuronales cuando se enfrentan a problemas que, aunque poseen causa y efecto, son muy complejos para derivar y expresar relaciones conductuales.

La intención es que la red neuronal aprenda a generalizar: es decir, habiendo aprendido a clasificar correctamente las muestras de aprendizaje, si se presenta una nueva muestra, la clasificara correctamente también, porque ha aprendido a reconocer las características que distinguen las diferentes clases de muestras.

4) Características importantes de muestras para entrenamiento

- Primera, las muestras deben reflejar todas las condiciones de entrada que están asociadas con la salida deseada, en cualquier combinación.
- Segunda, y mas critico, buenos contra-ejemplos, deben ser incluidos, Si se intenta sonar la campana una vez para una cena, y dos veces para evacuar el edificio, se habrá entrenado mejor al perro seguramente; de lo contrario cuando la campana suene dos veces el perro saldría disparado a la cocina en lugar de a la salida del edificio.
- Tercera, el viejo dicho: "no puedes enseñar a un perro viejo, nuevos trucos" es adecuado. Ya que los pesos en una red han sido ajustados para aprender un mapeo particular, es difícil de ajustarlos a aprender un mapeo nuevo o uno modificado. Si primero presentas puros ejemplos correctos y luego incorrectos, la red probablemente clasificara a todo como bueno. El problema de requerir redes neuronales para que retengan lo que han aprendido, y ser capaces de incorporar nuevos conocimientos cuando se requiera, se conoce como el dilema estabilidad-flexibilidad.
- Cuarta, se necesita concentrar esfuerzo donde la clasificación es mas difícil. Es mas fácil distinguir entre entradas muy diferentes que causan salidas muy diferentes que entre entradas ligeramente diferentes que causan salidas muy

diferentes. Si se debe tratar con ambos, la mayoría de los patrones deben de ser concentrados en las entradas difíciles de distinguir.

- Quinta, las redes neuronales son propensas a "sobre-aprender". Esto ocurre cuando la red reconoce una entrada para la que ha sido entrenada, pero falla en clasificar correctamente entradas no-vistas previamente, esto puede ocurrir si los mismos patrones en el procedimiento de entrenamiento son repetidamente presentados en un intento de hacer una clasificación mas precisa.

5) Problemas

Es difícil entender porque una red neuronal hace lo que hace. Aun si la red clasifica correctamente todas sus entradas, y aparenta generalizar nuevas entradas, no hay forma sencilla de predecir lo que pasara cuando una entrada no-vista se presenta. La única forma es haciendo la prueba. Las redes neuronales efectivamente, pueden ser clasificadas como una aproximación del enfoque de "caja negra" a la solución de problemas, adecuada cuando el diseñador necesita un problema resuelto, pero no le importa como, y puede tolerar decisiones pobres si ocurren.

6) Entrenamiento no supervisado

En los párrafos anteriores se trató con redes multi-nivel de alimentación hacia adelante. Estas son parte de las clases de redes neuronales que usan algoritmos de entrenamiento supervisados, donde el entrenador determina que salidas se esperan de las entradas de entrenamiento. Existen algoritmos de entrenamiento no-supervisados, donde la red aprende a clasificar entradas basada solamente en su similitud con otras entradas. Los algoritmos de aprendizaje no-supervisados son útiles cuando el usuario no sabe exactamente que clases se desean o esperan, también permiten a la red descubrir patrones entre los datos.

En el aprendizaje supervisado, hay un "profesor" que en la fase de aprendizaje "le dice" a la red como debe comportarse ("aprendizaje reforzado") o cual debe ser el comportamiento correcto ("completamente supervisado").

En el aprendizaje no supervisado: solo se muestran los datos, se hallan algunas de sus propiedades del conjunto de datos y aprende a reflejar esas propiedades en la salida. Cuales son esas propiedades exactamente, que la red aprenda a reconocer, depende del modelo particular de la red y del método de aprendizaje. Usualmente, la red aprende representaciones mas compactas de los datos.

7) Modelos de entrenamiento y topologías de red

1.- Entrenamiento no supervisado

- a. Matriz de aprendizaje.
- b. Aprendizaje reforzado.
- c. Memoria asociativa lineal.

- d. Memoria asociativa lineal optima.
- e. Memoria asociativa distribuida.
- f. Contrapropagacion.

(Todas son feed-forward) 2.- Entrenamiento supervisado

- 1. Adaline, Madaline.
- 2. Back-propagation.
- 3. Maquina de Cauchy.
- 4. Critica heurística adaptable.
- 5. Red neuronal delatora de tiempo.
- 6. Asociación premio-castigo.
- 7. Filtro avalancha.
- 8. Backpercolarion.
- 9. Artmap.
- 10. Red lógica adaptable.
- 11. Correlación cascada.
- 12. Filtro Kalman extendido.
- 13. Vector de aprendizaje cuantitativo.
- 14. Red neuronal probabilistica.
- 15. Red neuronal de regresión lineal.
- 16. Perceptron

IV) Aplicaciones de redes neuronales

1) Control de la eficiencia de una máquina

El comportamiento de la maquina de un carro es influenciada por un gran numero de parámetros: temperatura, mezcla aire-combustible, viscosidad del lubricante, etc. Rolls Royce ha usado redes neuronales para ajustar dinámicamente una maquina de acuerdo a su estado actual.

2) Reconocimiento de firmas

Cada persona tiene una firma distinta. Aunque cada vez que firmamos algo, la firma es ligeramente distinta, la forma en general es similar y puede ser reconocida por un experto humano. Una compañía ha creado una maquina que reconoce firmas con un gran nivel de precisión. Toma en cuenta la velocidad en que se firma, además de la presión, grosor y otros factores.

3) Reconocimiento de blancos mediante sonares

El propósito es de reconocer objetos "hostiles", en este caso, minas en el mar. Las minas deben ser distinguidas de entre otros objetos similares, pero inofensivos (como rocas). La red neuronal es provista con un gran numero de parámetros que son extraídos de la señal del sonar, cada uno de los cuales repercute en la forma del objeto. El conjunto de entrenamiento consiste de señales de minas y de rocas.

4) Predicción en la bolsa de valores

El triunfo (o no) en el uso real en la predicción de mercados es difícil de alcanzar desde que las compañías usando estas técnicas son comprensiblemente reacias a mostrar

información. Algunos creen que la proclama de usar redes neuronales es un truco de mercadeo.

5) Otras

- Reconocimiento óptico de caracteres
- Control de procesos industriales
- Aplicaciones de predicción del tiempo
- Decisiones sobre otorgamientos de préstamos, etc.
- Análisis de inversiones.
- Análisis de firmas.
- Monitoreo.
- Mercadotecnia.