**UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS**

Universidad del Perú, Decana de América

**FACULTAD DE SISTEMA E INFORMÁTICA**

**Escuela Profesional de Ingeniería de Software**



**Tema:** Estudio de la Red Neuronal ART

**Curso:** Sistemas Inteligentes

**Docente:** Chávez Herrera, Carlos Ernesto

**Integrantes:**

Terrazas García, Guillermo Moisés

Ramírez Vera, Carlos Augusto

Puitiza López, Anthony Brayan

Zuñiga Yamashita, Miguel Ángel

Olivares Chuquiure, Kevin

**Lima - Perú**

**Abril, 2018**

Contenido

[1. INTRODUCCIÓN 1](#_Toc511596574)

[1.1 ART (ADAPTIVE RESONANCE THEORY) 3](#_Toc511596575)

[1.2 HISTORIA DE LA RED NEURONAL ART 3](#_Toc511596576)

[2. DEFINICIÓN DE UNA RED ART 3](#_Toc511596577)

[2.1 CARACTERÍSTICAS 4](#_Toc511596578)

[2.2 CREADORES DE LA RED ART 4](#_Toc511596579)

[2.2.1 Stephen Grossberg 4](#_Toc511596580)

[2.2.2 Gail Carpenter 6](#_Toc511596581)

[3. TIPOS DE REDES ART 7](#_Toc511596582)

[3.1 ART-1: 7](#_Toc511596583)

[3.2 ART-2: 7](#_Toc511596584)

[3.3 ART-2ª: 7](#_Toc511596585)

[3.4 ART-3: 7](#_Toc511596586)

[3.5 ART DIFUSO (Fuzzy ART): 8](#_Toc511596587)

[3.6 ARTMAP (Fuzzy ARTMAP): 8](#_Toc511596588)

[4. ADAPTACIONES DE REDES ART 9](#_Toc511596589)

[5. ARQUITECTURA 9](#_Toc511596590)

[5.1 EL SUBSISTEMA DE ORIENTACIÓN 9](#_Toc511596591)

[5.2 EL SUBSISTEMA DE ATENCIÓN 9](#_Toc511596592)

[5.3 DILEMA ELASTICIDAD - PLASTICIDAD 9](#_Toc511596593)

[6. RECONOCIMIENTO 9](#_Toc511596594)

[7. ALGORITMO DE APRENDIZAJE 9](#_Toc511596595)

[8. LIMITACIONES DE LA RED ART 9](#_Toc511596596)

[9. LA RED ART-1 9](#_Toc511596597)

[9.1 ESQUEMA ART-1 9](#_Toc511596598)

[10. LA RED ART 2 9](#_Toc511596599)

[10.1 DESCRIPCION DE LA RED ART 2 9](#_Toc511596600)

[10.2 ARQUITECTURA DE ART 2 10](#_Toc511596601)

[10.3 PROCESAMIENTO EN F1 11](#_Toc511596602)

[10.4 PROCESAMIENTO EN F2 14](#_Toc511596603)

[10.5 ECUACIONES LTM 14](#_Toc511596604)

[10.6 SUBSISTEMA ORIENTADOR DE ART 2 15](#_Toc511596605)

[10.7 INICIALIZACIÓN LTM BOTTON-UP 17](#_Toc511596606)

[10.8 RESUMEN DEL PROCESAMIENTO DE ART2 18](#_Toc511596607)

[11. LA RED ART 3 21](#_Toc511596608)

[12. COMPARACION DE VENTAJAS Y DESVENTAJAS DE LA RED ART CON OTRAS REDES NEURONALES ARTIFICIALES 21](#_Toc511596609)

[13. APLICACIONES 21](#_Toc511596610)

[14. EJERCICIOS 21](#_Toc511596611)

[BIBLIOGRAFÍA 22](#_Toc511596612)

# INTRODUCCIÓN

El cerebro Humano ha sido considerado como el hombre como una maquina perfecta con capacidades inimaginables, que si se llegará a imitar alcanzaría grandes aplicaciones en ámbitos económicos, científicos y además de múltiples beneficios en entornos sociales. Los científicos en la necesidad de crear diferentes tipos de mecanismos para la simulación del aprendizaje y razonamiento humano han creado distintos métodos entre los cuales se encuentran Redes Neuronales Artificiales (RNA).

Durante la última década en el campo de la recuperación de la información se ha experimentado con diversas técnicas de Inteligencia Artificial (IA) basada en reglas y conocimiento. Estas técnicas en su tiempo parecían tener limitaciones y dificultades de aplicación, por lo que ya en el presente se ha comenzado a trabajar con las técnicas de IA más reciente, basadas en el aprendizaje inductivo: el aprendizaje simbólico, los algoritmos genéticos y las redes neuronales. (Chen, 1995)

Los primeros trabajos en computación neuronal se remontan a principios de los años 40. Un neurofísico llamado Warren McCulloch y un matemático llamado Walter Pitts propusieron, en base a sus estudios del sistema nervioso, un modelo de neurona formal implementada mediante circuitos eléctricos (McMulloch & PITTS, 1943). El entusiasmo que despertó el modelo neuronal impulsó la investigación en esta línea durante los años 50 y 60. En 1957 Frank Rosenblatt desarrolló el Perceptrón, un modelo de red que posee la capacidad de generalización, por lo que se ha utilizado hasta el día de hoy en diversas aplicaciones, generalmente en el reconocimiento de patrones. En 1959 Bernard Widrow y Marcial Hoff, de la Universidad de Stanford, desarrollaron el modelo ADALINE (ADAptative LINear Elements), primera RNA aplicada a un problema real (filtros de ruidos en líneas telefónicas).

En 1969 Marvin Minsky y Seymour Papert, del MIT, publicaron una obra en la que atacan al modelo neural y consideran que cualquier investigación en esa línea era estéril (Minsky & Papert, 1969). Debido a esta crítica los trabajos sobre RNA se detienen hasta un nuevo impulso durante los años 80. A pesar de esta pausa, varios investigadores siguieron trabajando en esa dirección durante los años 70. Tal es el caso del estadounidense James Anderson que desarrolla el modelo BSB (Brain-State-in-a-Box), o del finlandés Teuvo Kohonen que hace lo propio con uno basado en mapas auto-organizativos

A partir de 1982 el interés por la computación neuronal comenzó a cobrar fuerza nuevamente. El avance logrado en hardware y software, los avances metodológicos en torno a los algoritmos de aprendizaje para RNA, y las nuevas técnicas de inteligencia artificial, favorecieron este renacimiento. Ese mismo año se realiza la primera conferencia entre investigadores de computación neuronal de EEUU y Japón. En 1985 el Instituto Americano de Física establece la reunión anual *Neural Networks for Computing*. En 1987 el IEEE celebró la primera conferencia internacional sobre RNA. Ese mismo año se crea la Sociedad Internacional de Redes Neuronales (INNS).

Durante los años 90, las RNA comienzan a ser aplicadas a distintos campos del conocimiento: clasificación de patrones, robótica, visión artificial, procesamiento de señales, reconocimiento de escritura y habla, etc. En este contexto, las aplicaciones documentales no han podido quedar al margen de este proceso.

## ART (ADAPTIVE RESONANCE THEORY)

Es una teoría neuronal y cognitiva de como el cerebro aprende a categorizar, reconocer y predecir objetos y eventos en un mundo cambiante. ART aclara los procesos cerebrales a partir de los cuales experiencias conscientes emergen. Predice un enlace funcional entre los procesos de Conciencia, Aprendizaje, Expectativa, Atención, Resonancia y Sincronía (CLEARS), incluido la predicción de que "todos los estados conscientes son estados resonantes". Esta conexión aclara cómo en el cerebro la dinámica permite a un individuo, adaptarse de forma autónoma en tiempo real a un entorno cambiante.

## HISTORIA DE LA RED NEURONAL ART

Esta red fue desarrollada por los científicos americanos Stephen Grossberg y Gail Carpenter en 1987. Stephen Grossberg es un científico cognitivo, matemático, ingeniero biomédico y profesor en la universidad de Boston. Carpenter es el actual director del Departamento de Sistemas Cognitivos y Neuronales (CNS) Laboratorio de Tecnología de la Universidad de Boston. Grossberg y Carpenter desarrollaron la Teoría de Resonancia Adaptiva en respuesta a este dilema, en el que se plantean las siguientes cuestiones:

* **Plasticidad del aprendizaje:** permite a una red neuronal poder aprender nuevos patrones.
* **Estabilidad del aprendizaje:** permite a una red neuronal poder retener los patrones aprendidos.

# DEFINICIÓN DE UNA RED ART

La red neuronal se basa en la competencia y utiliza un modelo de aprendizaje no supervisado. Las redes de Teoría de Resonancia Adaptativa (ART, por sus siglas en inglés), como su nombre indica, siempre están abiertas a nuevos aprendizajes (adaptativos) sin perder los viejos patrones (resonancia). Básicamente, la red ART es un clasificador de vectores que acepta un vector de entrada y lo clasifica en una de las categorías dependiendo de cuál de los patrones almacenados se asemeja más.

## CARACTERÍSTICAS

* El aprendizaje se produce mediante un mecanismo de realimentación creado por la competencia entre las neuronas de la capa de salida y la capa de entrada de la red.
* Tipo [no supervisado](https://es.wikipedia.org/wiki/Aprendizaje_no_supervisado), es decir, no hay un conocimiento a priori, aunque existe una modalidad supervisada.
* Se crea una propia clasificación de lo que se aprende

## CREADORES DE LA RED ART

**ART** son las siglas en inglés de Teoría de la Resonancia Adaptativa (*Adaptive Resonance Theory*), desarrollada por [Stephen Grossberg](https://es.wikipedia.org/w/index.php?title=Stephen_Grossberg&action=edit&redlink=1) y [Gail Carpenter](https://es.wikipedia.org/w/index.php?title=Gail_Carpenter&action=edit&redlink=1) en 1986.

### [Stephen Grossberg](https://es.wikipedia.org/w/index.php?title=Stephen_Grossberg&action=edit&redlink=1)

Científico cognitivo, neurocientífico, matemático, ingeniero biomédico y tecnólogo neuromórfico. Profesor de sistemas cognitivos y neuronales, matemáticas, psicología e ingeniería biomédica en la Universidad de Boston.



Figura 1: Grossberg in July, 2016. (Fuente https://en.wikipedia.org/wiki/Stephen\_Grossberg)

Grossberg es uno de los fundadores de los campos de la neurociencia computacional, la ciencia cognitiva conexionista y la tecnología neuromórfica. Su trabajo se centra en los principios y mecanismos de diseño que permiten el comportamiento de los individuos, o máquinas, para adaptarse de forma autónoma en tiempo real a los desafíos ambientales inesperados. Esta investigación ha incluido modelos neuronales de visión y procesamiento de imágenes; aprendizaje de objetos, escenas y eventos, reconocimiento de patrones y búsqueda; audición, habla y lenguaje; procesamiento y planificación de la información cognitiva; aprendizaje de refuerzo e interacciones cognitivo-emocionales; navegación autónoma; control adaptativo sensor-motor y robótica; neurodinámica auto organizada; y trastornos mentales. Grossberg también colabora con experimentales para diseñar experimentos que prueban predicciones teóricas y rellenan huecos conceptualmente importantes en la literatura experimental, lleva a cabo análisis de la dinámica matemática de sistemas neuronales y transfiere neuronas biológicas modelos para aplicaciones en ingeniería y tecnología. Ha publicado diecisiete libros o números especiales en revistas, más de 500 artículos de investigación y tiene siete patentes.

Grossberg ganó el primer IEEE *Neural Network Pioneer Award*, el INNS Leadership Award de 1992, el 1992 *Boston Computer Society Thinking Technology Award*, el 2000 *Information Science Award de la Association for Intelligent Machinery*, el 2002 Charles River Laboratories, premio de la *Society for Behavioral Toxicology* y el Premio INNS Helmholtz de 2003. Es miembro de 1990 de *Memory Disorders Research Society*, miembro de 1994 de la Asociación Americana de Psicología, miembro de la Sociedad de Psicólogos Experimentales de 1996, miembro de la Sociedad Americana de Psicología en 2002, miembro de IEEE en 2005, miembro inaugural de 2008 de la Asociación Estadounidense de Investigación Educativa, y un Miembro de INNS 2011. Grossberg recibió el Premio Norman Anderson *Lifetime Achievement* 2015 de la Sociedad de Psicólogos Experimentales "por su investigación teórica pionera sobre cómo los cerebros dan lugar a las mentes y sus contribuciones fundamentales a la neurociencia computacional y cognitiva conexionista ciencia ". (GROSSBERG, 2015)

Recibió el Premio Frank Rosenblatt del Instituto de Ingenieros Eléctricos y Electrónicos 2017 (IEEE) con la siguiente cita: "Por sus contribuciones para comprender la cognición y el comportamiento del cerebro y su emulación por tecnología".

### Gail Carpenter

Científica cognitiva, neurocientífica y matemática. Profesora de sistemas cognitivos y neuronales y de matemáticas en la Universidad de Boston y directora de su laboratorio del Departamento de Sistemas Cognitivos.



Figura 2 : Gail Carpenter, September 2017 (Fuente: http://www.bu.edu/math/people/faculty/mathematical-biology-and-neuroscience/carpenter/)

Carpenter es la primera mujer en recibir el IEEE Neural Networks Pioneer Award. Ha sido elegida para mandatos sucesivos de tres años en la Junta de Gobernadores de la Sociedad Internacional de Redes Neuronales (INNS) desde su fundación, en 1987, y recibió el Premio INNS Gabor. También ha servido como miembro electo del Consejo de la American *Mathematical Society*, y es miembro fundador de la Asociación de Mujeres en Matemáticas.

# TIPOS DE REDES ART

Hay dos grandes ramas de redes ART: las ART1 que trabajan con vectores de información binaria y las redes ART2, que procesan información analógica. Sin embargo, con el desarrollo de las redes neuronales, de la computación y de las nuevas aplicaciones han surgido varios subgrupos de redes ART:

## ART-1:

Es el tipo más sencillo de red ART, sólo acepta entradas binarias (Carpenter & Grossberg, 2003). Trabaja con un operador duro, el cual hace posible las salidas binarias. Aplicaciones: Reconocimiento de letras y figuras con patrones de colores en blanco y negro. (Grossberg, 1987).

## ART-2:

Aumenta las capacidades de la red al soportar estos valores de entrada continuos (Carpenter & Grossberg, S, 1987).

## ART-2ª:

Es una versión más rápida del algoritmo de redes ART2. Es una forma simplificada de ART-2 con un tiempo de ejecución drásticamente acelerado, y con resultados cualitativos que rara vez son inferiores a los obtenidos por una red ART-2 completa. (Carpenter, Grossberg, & Rosen, 1991).

## ART-3:

Extensión de red ART que incorpora transmisores químicos para controlar el proceso de búsqueda de categorías dentro de la red. ​Se basa en ART-2 mediante la simulación de rudimentarios [neurotransmisores](https://es.wikipedia.org/wiki/Neurotransmisores) y regulación de  [actividad sináptica](https://es.wikipedia.org/wiki/Sinapsis) mediante la incorporación de simulaciones de concentraciones de iones de sodio (Na +) y calcio (Ca2+) en las ecuaciones del sistema, lo que origina más realismo en la simulación de las condiciones fisiológicas en las que se disparan las neuronas biológicas. (Carpenter & Grossberg, 1990)

## ART DIFUSO (Fuzzy ART):

Aplica la lógica difusa en el reconocimiento de patrones de ART, aumentando así la generalización. Una opcional y muy útil característica de ART difuso es el complemento de código, una forma de incorporar la ausencia de elementos en las clasificaciones de patrones que sirve para prevenir la creación de categorías ineficientes e innecesarias. (Carpenter, Grossberg, & Rosen, 1991).

## ARTMAP (Fuzzy ARTMAP):

Es una versión supervisada de ART que puede aprender mapas de patrones binarios. También conocido como ART Predictivo, combina de unidades de ART-1 o ART-2 ligeramente modificadas formando una estructura de aprendizaje supervisado, donde la primera unidad tiene los datos de entrada y la segunda unidad toma la salida de datos correctos, se utiliza para posibilitar el mínimo ajuste del parámetro de vigilancia en la primera unidad con el fin de obtener una clasificación correcta. (Carpenter G. , Grossberg, Markuzon, Reynolds, & Rosen, 1992).

# ADAPTACIONES DE REDES ART

# ARQUITECTURA

## EL SUBSISTEMA DE ORIENTACIÓN

## EL SUBSISTEMA DE ATENCIÓN

## DILEMA ELASTICIDAD - PLASTICIDAD

# RECONOCIMIENTO

# ALGORITMO DE APRENDIZAJE

# LIMITACIONES DE LA RED ART

# LA RED ART-1

## ESQUEMA ART-1

# LA RED ART 2

## DESCRIPCION DE LA RED ART 2

ART2 difiere de ART1 solamente en la naturaleza de los patrones de entrada: ART2 acepta componentes de vectores análogos (escala de gris), así como también componentes binarios. Esta capacidad representa un realce al sistema.

La diferencia entre ART1 y ART2 se basan en diferencias de arquitectura que le da a ART2 su estabilidad para enfrentarse a patrones análogos. Estas diferencias a veces son más complejas, y a veces menos complejas que las estructuras ART1.

ART2 debe enfrentarse con complicaciones adicionales. Por ejemplo, ART2 debe poder reconocer la similitud subyacente de patrones idénticos sobrexpuestos en fondos constantes teniendo diferentes niveles. Comparado en un sentido común, dos patrones similares deben aparecer enteramente diferentes cuando en realidad, se debe estar clasificado el mismo patrón.

El precio de esta capacidad adicional es primeramente un incremento en la complejidad del procesamiento del nivel de F1. El nivel en ART1 consiste en diferentes subniveles y varios sistemas de control de ganancia. El procesamiento en F2 es igual tanto para ART2 como para ART1. Como una compensación parcial para la complejidad añadida, las ecuaciones LTM son un poco más simples para ART2 que para ART1. (Carpenter & Grossberg, S, 1987)

## ARQUITECTURA DE ART 2

Como se mencionó en la introducción de esta sección, ART2 tiene un parecido superficial a ART1. los dos tienen un subsistema de atención y un subsistema orientador. El subsistema de atención para cada arquitectura consiste en dos capas de elementos de procesamiento, F1 y F2, y, un sistema de control de ganancia. El subsistema orientador para cada red realiza la función idéntica. Además, las ecuaciones diferenciales básicas que gobiernan las actividades de los elementos de procesamientos individuales son los mismos en los dos casos. Para enfrentarse exitosamente con patrones análogos en el ART2, se tuvo que separar la capa F1 en un número de subcapas conteniendo las conexiones Feedforward y Feedback. La figura 3 muestra la estructura resultante.

Toda la estructura de la red ART2 es similar que la de ART1 (Figura 3). la capa F1 ha sido dividida en 6 subcapas w, x, u, v, p y q. Cada nodo marcado con G es una unidad de control de ganancia que envía una señal de inhibición no específica a cada unidad en la capa que este alimenta. Todas las subcapas en F1, así como en la capa r del subsistema orientador, tiene el mismo número de unidades. Las subcapas individuales en F1 están conectadas de unidad a unidad; esto es, que las capas no están totalmente interconectadas, con excepción de las conexiones Botton-Up a F2 y las conexiones Top-Down de F2.

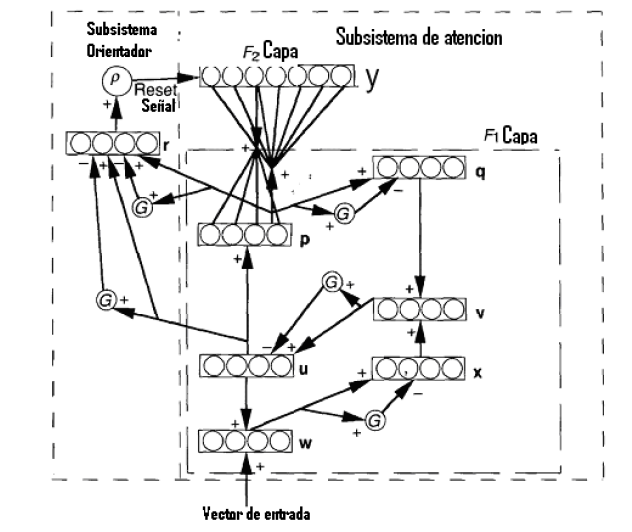


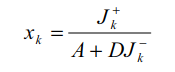
Figura 3. Arquitectura de ART 2(Tomado de bibliografía 3. Capítulo 8, Adaptive Resonance Theory)

## PROCESAMIENTO EN F1

La actividad de cada unidad en cada subcapa de F1 es gobernada por una ecuación de la forma

 (30)

Donde A, B, C, D son constantes. La ecuación 30 es casi idéntica a la ecuación 1 de ART1 la única diferencia es de un factor multiplicativo en el primer término en el lado derecho de la ecuación 30. Para el modelo de ART2 presentado, se debe poner B y C idénticamente igual a cero. Igual que en ART1, Jk+ y Jk- representa la excitación de la red y los factores inhibidores respectivamente. Igual que en ART1 se debe estar interesado en la solución asintótica, entonces

 (31)

Los valores de las cantidades individuales en la ecuación 31 varían de acuerdo a la subcapa en que se esté considerando. Por conveniencia, se realizó la tabla 7, el cual muestra todas las cantidades apropiadas para cada subcapa F1, así también, en capa r del subsistema orientador. Basada en la tabla 7, las actividades en cada una de las 6 subcapas pueden ser reducidas en las siguientes ecuaciones.

 (32)

 (33)

 (34)

 (35)

 (36)

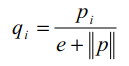
 (37)

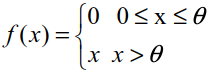
Tabla 1  
 Cantidades apropiadas para cada subcapa F1, la capa r del subsistema orientador

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| CAPA | A | D | J + | J − |
| W | 1 | 1 | Ii + aui | 0 |
| X | e | 1 | wi | ||w|| |
| u | e | 1 | vi | ||v|| |
| v | 1 | 1 | f(xi) + bf(qi) | 0 |
| p | 1 | 1 | ui+ | 0 |
| q | e | 1 | ρ i | ||p|| |
| r | e | 1 | u + cp | ||u||+ c||p|| |

Tomado de bibliografía [3]

Los factores en la ecuación 31 para cada subcapa F1 y la capa r. Ii es el iteavo componente del vector de entrada. Los parámetros a, b, c y e son constantes cuyos valores serán discutidos en esta investigación. Yj es la actividad de la jenésima unidad en la capa F2 y g(y) es la función de salida de F2. La función f(x) es descrita en esta investigación.

Se Explicará la capa **r** del subsistema orientador más adelante. El parámetro **e** se le coloca un numero positivo considerablemente menor que uno. Este tiene el efecto de mantener las activaciones finitas cuando no hay entrada presente en el sistema. No se requiere la presencia de e para esta discusión por eso se colocará e = 0 para lo que resta de ART2. Las tres unidades de control de ganancia en F1 inhibe no específicamente la subcapa x, u y q. La señal inhibidora es igual a la magnitud del vector de entrada a esas capas. El efecto es que las actividades de estas tres capas son normalizadas por unidad por las señales de control de ganancia. La forma de la función f(x), determina la naturaleza de realce de contraste que pasa en F1.

 (38)

Donde θ > 1, se utilizará a θ = 0.2

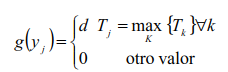
## PROCESAMIENTO EN F2

El procesamiento de ART2 en F2 es igual que en ART1. Las entradas Botton-Up son calculadas como en ART1:

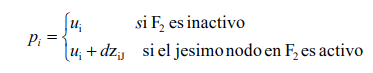
 (39)

La competencia en F2 resulta en un realce de contraste donde se escoge un solo nodo ganador.

La función de salida de F2 está dada por

 (40)

Esta ecuación presume que el ajuste de {TK} incluye solamente los nodos que no han sido reseteados recientemente por el subsistema orientador. Ahora se puede reescribir la ecuación para el procesamiento en la capa p de F1 como (Eq 36).

 (41)

## ECUACIONES LTM

Las ecuaciones LTM en ART2 son significativamente menos complejas que las presentadas en ART1. Las ecuaciones Botton-Up y Top-Down tienen la misma forma:

 (42)

Para los pesos Botton-Up de vi en F1 a vj en F2, y

 (43)

Para los pesos Top-Down de vj en F2 a vi en F1. Si vj es el nodo F2 ganador, se puede usar la ecuación 40 en la ecuación 41, 42 y 43 para mostrar que



Y similarmente



Con los otros ziJ = zJi = 0 para j ≠ J. Se debe estar interesado en el caso de rápido aprendizaje, entonces se puede resolver para los valores de equilibrio de los pesos

 (44)

Se asume que 0 < d < 1.

Se debe posponer la discusión para los valores iníciales para los pesos hasta después de la discusión del subsistema orientador.

## SUBSISTEMA ORIENTADOR DE ART 2

De la tabla 1 y la ecuación 31 se puede construir la ecuación para las actividades de los nodos en la capa **r** del subsistema orientador, como **e** es cero

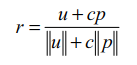
 (45)

La condición de reset es.

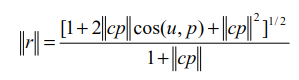
 (46)

Donde p es el parámetro de vigilancia.

Nótese que dos subcapas de F1 p y u participan en el proceso de concordancia. Así como los pesos Top-Down cambian en la subcapa p durante el aprendizaje, la actividad de las unidades de la subcapa p también cambian. La capa u se mantiene estable durante este proceso, entonces incluyendo a esta en proceso de aprendizaje previenen que ocurra el reset mientras el aprendizaje de un nuevo patrón se está llevando a cabo. Se puede reescribir la ecuación 45 en la forma de vector como.



Entonces, de , entonces se puede escribir

 (47)

Donde cos (u, p) es el cos del ángulo entre u y p. primero note que, si u y p son paralelos, entonces la ecuación 47 se reduce a ||r|| = 1, y allí no existirá reset. Mientras no exista salida de 2, la ecuación 36 muestra que u = p, y no existiría reset para este caso.

Suponga ahora que F2 tiene una salida de alguna unidad ganadora, y que los patrones de entrada necesitan ser aprendidos o codificados por la unidad F2. No se quiere un reset en este caso. De la ecuación 36 se tiene que p = u + dzj , donde la Jesima unidad en F2 es la ganadora y zJ = (ziJ , ziJ ,..., zMJ )t . si se inicializa todos los pesos Top-Down zij en cero, entonces la salida inicial de F2 no va a tener efecto en el valor de p; p se va a mantener igual a u.

Durante el proceso de aprendizaje, zJ se convierte paralelo a u de acuerdo a la ecuación 44. Además, p también se convierte paralelo a u, y otra vez ||r|| = 1 y no hay reset.

Así como ART1, una suficiente diferencia entre el vector de entrada Botton-Up y la plantilla Top-Down da resultando un reset. En ART2, el patrón Botton-Up es tomado del subnivel u de F1 y la plantilla Top-Down es tomado en p. Ya se vio que los pesos Top-Down deben estar inicializados en cero. La inicialización de los pesos Botton-Up se verá a continuación.

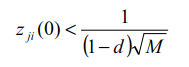
## INICIALIZACIÓN LTM BOTTON-UP

Se ha discutido la modificación de los rastros LTM, o pesos en el caso de rápido aprendizaje. Vamos a examinar el comportamiento dinámico de los pesos Botton-Up durante un proceso de aprendizaje. Asumimos que un nodo particular de F2 ha previamente codificado un vector de entrada tal que zji = ui / (1 − d), y, por lo tanto, ||zJ||= [ ||u||/ (1 − d)]− [1/1 − d], donde zj es el vector de los pesos Botton-Up en la Jesimo nodo F2. Suponga que el mismo nodo gana por un patrón de entrada un poco diferente, uno el cual el grado de diferencia no es suficiente para causar un reset. Entonces, los pesos Botton-Up van a ser decodificados para concordar el nuevo vector de entrada. Durante este proceso de recodificación dinámico, ||zJ|| puede disminuir antes de regresar al valor [1/1 − d]. Durante este periodo de disminución, ||r|| va a disminuir también. Si otros nodos han tenido sus valores de pesos inicializados tal que ||zJ (0)|| > 1/1−d, entonces la red debe cambiar los ganadores en el medio del proceso de aprendizaje.

Se debe, por lo tanto, inicializar los vectores de peso Botton-Up como:



Se puede realizar esta inicialización, poniendo los pesos a números pequeños aleatorios. Alternativamente, podemos utilizar la inicialización

 (48)

Este esquema anterior tiene una apariencia de una inicialización uniforme. Además, si se usa la igualdad, entonces los valores iníciales serán tan largos como sea posible. Haciendo los valores iníciales tan largos perjudica a la red hacia nodos sin compromisos. Aunque el parámetro de vigilancia este muy bajo para causar un reset, de otra manera la red va a escoger un nodo sin compromiso en vez de un nodo que tenga una mala concordancia. Este mecanismo ayuda a estabilizar la red en contra de la recodificación constante.

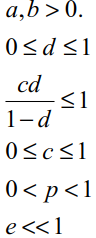
Similares argumentos llevan a un contraste en los parámetros c, d.

 (49)

Mientras el radio se acerca a uno, la red se vuelve más sensitiva a discordancias por que el valor de ||r|| disminuye a un valor pequeño, y las otras cosas siguen igual.

## RESUMEN DEL PROCESAMIENTO DE ART2

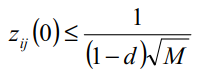
Se considera solamente las soluciones asintóticas para las ecuaciones dinámicas, y el modo de rápido aprendizaje. M va a ser el número de unidades en cada subcapa F1, y N será el número de unidades en F2. Los paramentos son escogidos de acuerdo a las siguientes restricciones



Los pesos Top-Down son inicializados en cero



Los pesos Botton-Up son inicializados de acuerdo a



Ahora estamos listos para procesar los datos

1. Inicializar todas las salidas de las capas y subcapas en vectores cero y establecer un ciclo de conteo inicializado a un valor de uno.
2. Aplicar un patrón de entrada I a la capa w de F1. La salida de esta capa es.



1. Propagar hacia la subcapa x.



1. Propagar hacia la subcapa v.



Note que el segundo término es cero a través del primer paso, así como es cero para ese tiempo.

1. Propagar hacia la subcapa u



1. Propagar hacia la subcapa p



Donde el Jesimo nodo en F2 es el ganador de la competencia en esa capa. Si F2 es inactivó, pi = ui. Similarmente, si la red esta aun en su configuración inicial, pi = ui porque zij (0) =0

1. Propagar hacia la subcapa q.



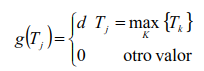
1. Repetir desde el segundo paso hasta el séptimo tantas veces sea necesario hasta estabilizar los valores en F1
2. Calcular la salida de capa r



1. Determinar si una condición de reset es indicada. Si p (e + ||r||) > 1, entonces enviar una señal de reset a F2. Marcar cualquier nodo F2 activo como inelegible para la competición, resetear el contador de ciclos a uno, y regresar al paso dos. Si no hay reset, y el contador de ciclos es uno, incrementar el contador de ciclos y continuar el paso 11. Si no hay reset, y el contador de ciclos es más grande que uno, entonces se saltará al paso 14, en donde una resonancia ha sido establecida.
2. Propagar la salida de la subcapa p a la capa F2. Calcular las entradas de red para F2



1. Solamente el nodo F2 ganador tiene una salida diferente a cero.



Todos los nodos marcados como inelegible por las señales previas de reset no participan en la competencia.

1. Repetir pasos 6 hasta el 10
2. Unificar los pesos Botton-Up en la unidad F2 ganadora.



1. Modificar los pesos Top-Down en la unidad F2 ganadora.



1. Remover el vector de entrada. Restaurar todas las unidades F2 inactivas.

Regresar al paso uno con un nuevo patrón.

# LA RED ART 3

# COMPARACION DE VENTAJAS Y DESVENTAJAS DE LA RED ART CON OTRAS REDES NEURONALES ARTIFICIALES

# APLICACIONES

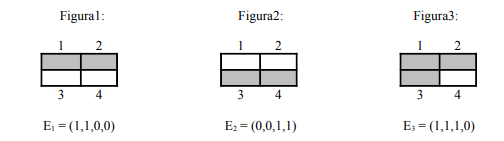
La utilización de la red ART suele estar relacionada con tareas de reconocimiento de patrones, ya que se trata de una red especialmente útil en aquellas aplicaciones que requieran el establecimiento automático de categorías para la clasificación de datos binarios (ART1) y continuos (ART2). También se ha utilizado para el diseño de sistemas de control y diagnóstico adaptativo.

Otras áreas de aplicación tecnológicas incluyen diseño industrial, el control de robot móviles, reconocimiento de rostro, reconocimiento de objetivos, diagnósticos médicos, análisis de electrocardiograma, verificación de firmas, reconocimiento visual de objetos 3D, reconocimiento sonar y radar. Debido al carácter de funcionamiento autónomo, sin supervisor, esta red ha sido también utilizada para modelar procesos biológicos y así poder estudiar y predecir su comportamiento.

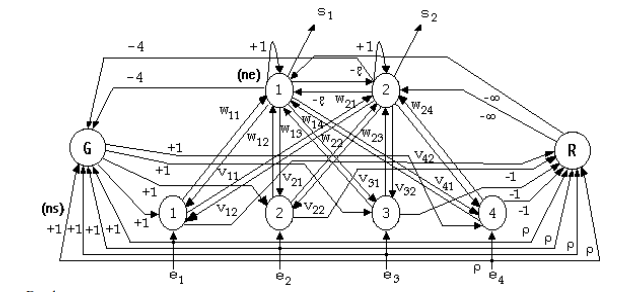
# EJERCICIOS

En este ejercicio utilizaremos ART1 con 4 neuronas en la capa de entrada y 2 neuronas en la capa de salida, está red por lo ya sabemos clasifica entradas o vectores binarios, el número de clasificaciones que puede hacer esta red está dada por el número de neuronas de salida, en este caso tenemos 2 neuronas de salida, lo que significa que habrá y reconocerá 2 categorías y sus variantes

Se utilizarán los siguientes vectores de entrada como patrones representados en forma de figuras:



Como se observa los píxeles negros son representados mediante el valor binario 1, y los blancos con el valor 0. Nuestra red quedaría con la siguiente estructura:



Donde:

* ρ es el parámetro de vigilancia, que tendrá el valor de 0.5
* ε para asegurar la convergencia de la red esté suele tener un valor de 1/M, donde M es el número de neuronas de salida, esto es 1/M=1/2=0.5

Los pesos de las conexiones feedforward (wji) y feedback(vij) se inicializan con las siguientes ecuaciones:

(1)

(2)

Sustituyendo la ecuación (1) tenemos:

Donde N es el número de neuronas de entrada. Las matrices que representan estos pesos iniciales serían:

Ya inicializada la red, el vector de entrada es aplicado:

(3)

Utilizando la ecuación anterior, obtenemos los siguientes valores:

Como está es la primera información de entrada y los pesos wji son todos iguales, no hay ninguna neurona vencedora en la competición, por lo cual se almacena en la primera neurona de la capa de salida, con la siguiente ecuación:

(4)

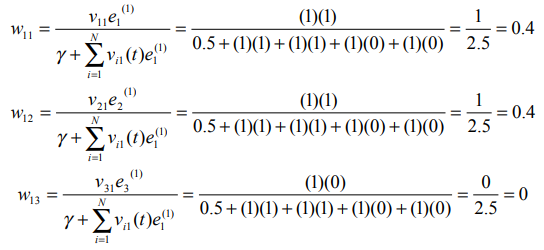
Sustituyendo la ecuación (4):

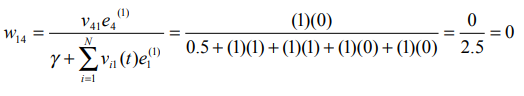
Así, los nuevos valores de V son:

Así, para obtener los pesos de W se obtienen con la siguiente ecuación:



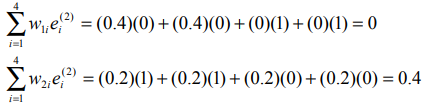
Sustituyendo la ecuación (5):





Obteniéndose los nuevos valores de W:

Si a continuación se introduce el siguiente vector de entrada E2 se producirá la competición en las neuronas de salida, sustituyendo la ecuación (3) con los siguientes resultados:



Por lo tanto resulta vencedora la segunda neurona de la capa de salida por lo que al sustituir la ecuación (3) y (4) el resultado es el siguiente:

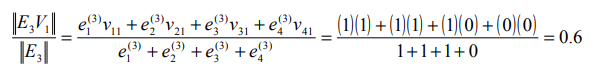
Por último, se presenta a la entrada el vector E3=(1,1,1,0) produciéndose la competición con la ecuación (3) con los siguientes resultados:



Como se observa la neurona vencedora es la primera de la capa de salida, por lo utilizaremos la siguiente ecuación de relación de semejanza:



Sustituyendo la ecuación anterior tenemos:



En este punto se deberá determinar si es suficiente la semejanza como para considerar que E3= (1,1,1,0) es de la misma clase que E1= (1,1,1,0), para lo cual habrá que tener en consideración el parámetro de vigilancia. El cual es de 0.5, entonces 0.6>0.5, por lo que se consideraría de la misma clase y se procedería a modificar los pesos de V y W de la primera neurona, utilizando la ecuación (4) y la ecuación (5) para el ajuste de los pesos. Por lo que los nuevos valores serán:

# BIBLIOGRAFÍA

Carpenter, G., & Grossberg, S. (1987). ART 2: Self-organization of stable category recognition codes for analog input patterns. *Applied Optics*, 4919-4930.

Carpenter, G., & Grossberg. (2003). *The Handbook of Brain Theory and Neural Networks* (Vol. II). Cambridge, MA: Michael A. Arbib.

Carpenter, G., & Grossberg, S. (1990). *Hierarchical search using chemical transmitters in self-organizing pattern recognition architectures*, 129-152.

Carpenter, G., Grossberg, S., & Rosen, D. B. (1991). An adaptive resonance algorithm for rapid category learning and recognition. *Neural Networks (Publication)*, 493-504.

Carpenter, G., Grossberg, S., & Rosen, D. B. (1991). Fast stable learning and categorization of analog patterns by an adaptive resonance system. *Neural Networks (Publication)*, 759-771.

Carpenter, G., Grossberg, S., Markuzon, N., Reynolds, J. H., & Rosen, D. B. (1992). A neural network architecture for incremental supervised learning of analog multidimensional maps. *IEEE Transactions on Neural Networks*, 698-713.

Chen, H. (1995). *Machine learning for information retrieval: neural networks, symbolic learning, and genetic algorithms* (Vol. 46). California: Journal of the American Society for Information Science.

Grossberg. (1987). *Competitive learning: From interactive activation to adaptive resonance.*

GROSSBERG, S. (2015). The 2015 Norman Anderson Lifetime Achievement Award is presented to. *SED Society of Experimental Psychologists*. Retrieved from http://www.sepsych.org/awards.php

Grossberg, S. (2015, Mayo 15). *Youtube*. Retrieved from Linking Mind and Brain: Steve Grossberg SEP Lifetime Achievement Award Lecture: https://www.youtube.com/watch?v=9n5AnvFur7I

McMulloch, W., & PITTS, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 115-133.

Minsky, M., & Papert, S. (1969). *Perceptrons.* Cambrige: MIT Press.