

## ANEXO B REDES NEURONALES ARTIFICIALES BASADAS EN ART

La ART fue desarrollada para dar respuesta al dilema de la estabilidad (capacidad de retener patrones aprendidos) y la plasticidad (capacidad de poder aprender nuevos patrones) del aprendizaje. Conseguir que un modelo de red neuronal artificial sea capaz de resolver uno de estos problemas es sencillo, el reto está en conseguir un modelo que sea capaz de dar respuesta a ambos. Las redes más conocidas, tales como el Perceptrón multicapa o el Adaline (Widrow & Lehr, 1998), son capaces de aprender cómo han de responder ante unos patrones de entrada pero, una vez entrenados, el intentar que aprendan nuevos patrones puede suponer el "olvido" de lo aprendido previamente.

Este tipo de clasificadores se caracterizan fundamentalmente por el método de aprendizaje que utilizan. El aprendizaje se produce mediante un mecanismo de realimentación creado por la interacción entre las neuronas de la capa de salida y la capa de entrada. Además, es un aprendizaje no supervisado, lo que aumenta considerablemente la automatización del proceso.

Respecto a la arquitectura de la red, se divide en tres capas: en primer lugar existe una capa de entrada donde no se realiza ningún preprocesado de los datos de entrada. Tras esta capa de entrada existe una (o varias) capa oculta, y tras la capa oculta se encuentra la capa de salida. Debe señalarse que la capa de entrada y la capa oculta tienen el mismo número de neuronas. Cada neurona de la capa de entrada es entrada de una única neurona en la capa oculta.

El funcionamiento de las redes ART está basado en el cálculo de dos tipos de pesos, los pesos  $W$ , que son pesos *feedforward* entre la capa oculta y la capa de salida, y los pesos  $T$ , que son los pesos *feedback* entre la capa de salida y la capa oculta. Los pesos *feedforward* ( $W$ ) tienen el mismo valor que los pesos *feedback* ( $T$ ), pero normalizados:

$$W = \frac{T}{\sum T + \gamma} \quad (\text{Ecuación B.1})$$

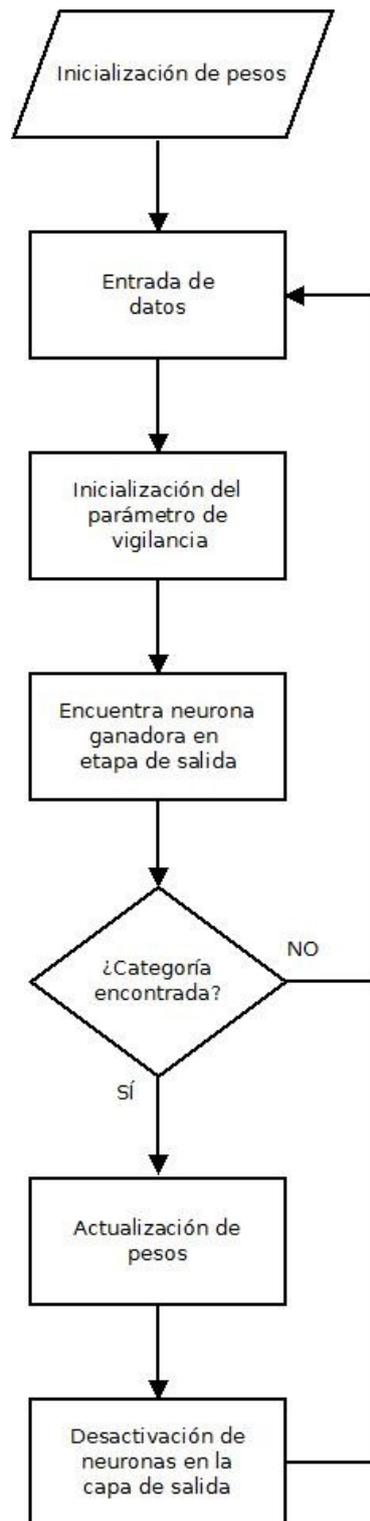
donde  $\gamma$  es un parámetro que toma generalmente el valor 0.5, menos en la inicialización de los pesos, que se recomienda que valga 1. El funcionamiento de un modelo ART se divide en cuatro fases:

- **Fase de inicialización.** En esta fase se inicializan los parámetros de la red neuronal y se establecen las señales de control. Los pesos son inicializados de la siguiente manera:

$W_{ij} = 1/(1 + N)$  para las conexiones hacia delante, siendo  $N$  el número de entradas a la red neuronal, y  $T_{ij} = 1$  para las conexiones hacia atrás. Mediante las señales de control (ganancia y reinicio) se controla el flujo de datos a través de la red neuronal y se pasa a las distintas fases.

- **Fase de reconocimiento.** En esta fase se efectúa una operación con los datos de entrada y los pesos  $W$  asociados a cada neurona de la capa de salida. El resultado de esta operación debe indicar qué clase tiene mayor prioridad para ver si los datos de entrada entran en resonancia con ella. A modo de ejemplo, podría calcularse la distancia euclídea entre los datos de entrada y los pesos  $W$ , y la clase ganadora sería aquella cuyo  $W$  estuviese más próximo a los datos de entrada y, por lo tanto, sería la primera clase a la que se le intentaría asociar dicho patrón.
- **Fase de comparación.** El vector de entrada y el vector producido por la capa de salida son comparados en la capa de entrada, y el resultado obtenido se envía al control de reinicio. El propósito de esta fase es conseguir una medida de similitud entre el vector de entrada y el vector prototipo que obtenemos a la salida.
- **Fase de búsqueda.** Si la neurona ganadora no representa la categoría del vector de entrada, esta neurona se desactiva y se empieza la búsqueda por otras categorías que ya posee la red. Se repiten por lo tanto los pasos anteriores hasta que se encuentre una nueva neurona ganadora que represente la categoría del vector de entrada. En el caso de que no se encontrara ninguna neurona ganadora para representar la categoría del vector de entrada, se llegaría a una situación de saturación de la red que podría solucionarse ampliando el número de neuronas de la red neuronal de forma dinámica.

En la Ilustración B.1 podemos observar un diagrama de flujo que representa el funcionamiento básico de una red ART.



**Ilustración B.1:** Diagrama de flujo de una red neuronal ART básica.

Un sistema ART básico sigue un modelo de aprendizaje no supervisado. Por regla general, consta de un campo de comparación y un campo de reconocimiento, que a su vez se compone de un parámetro de vigilancia y de un módulo de reinicialización. El parámetro de vigilancia tiene una influencia considerable en el sistema: un mayor valor de este parámetro

produce recuerdos muy detallados, mientras que valores más pequeños de dicho parámetro producirán recuerdos más generales.

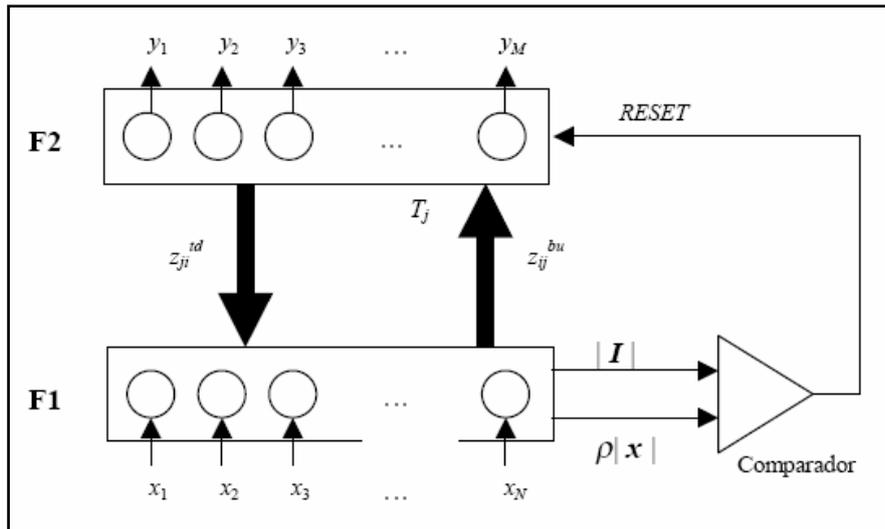
El campo de comparación toma un vector de entrada (una matriz bidimensional de valores) y transfiere su mejor coincidencia al campo de reconocimiento. Su mejor coincidencia estará en aquella neurona cuyo conjunto de pesos (vector de peso) se acerque más al vector de entrada. Cada neurona del campo de reconocimiento emite una señal que es proporcional a la calidad de coincidencia de dicha neurona con el vector de entrada, para cada una de las neuronas del campo de reconocimiento provocando una inhibición de su valor de salida. De esta manera el campo de reconocimiento exhibe una inhibición lateral, permitiendo que cada neurona en él represente una categoría en la que se clasifican los vectores de entrada. Después de que el vector de entrada es clasificado, el módulo de reinicialización compara la intensidad de la coincidencia encontrada por el campo de reconocimiento con el parámetro de vigilancia. Si el umbral de la vigilancia se cumple, se inicia el entrenamiento. De lo contrario, si el nivel de coincidencia no cumple con el parámetro de vigilancia, la neurona de reconocimiento disparada se inhibe hasta que un vector de entrada se aplique el nuevo.

El entrenamiento se inicia sólo al final del procedimiento de búsqueda, en el cual las neuronas de reconocimiento son desactivadas una a una por la función de reinicio hasta que el parámetro de vigilancia queda satisfecho con una coincidencia de reconocimiento. Si ninguna coincidencia encontrada por las neuronas de reconocimiento supera el umbral de vigilancia, una neurona no comprometida se ajusta para que concuerde con el vector de entrada.

En nuestro caso hemos utilizado un tipo de red ART denominado *fuzzy ARTMAP*. Para entender su funcionamiento, haremos un breve repaso de las arquitecturas en las que se basa.

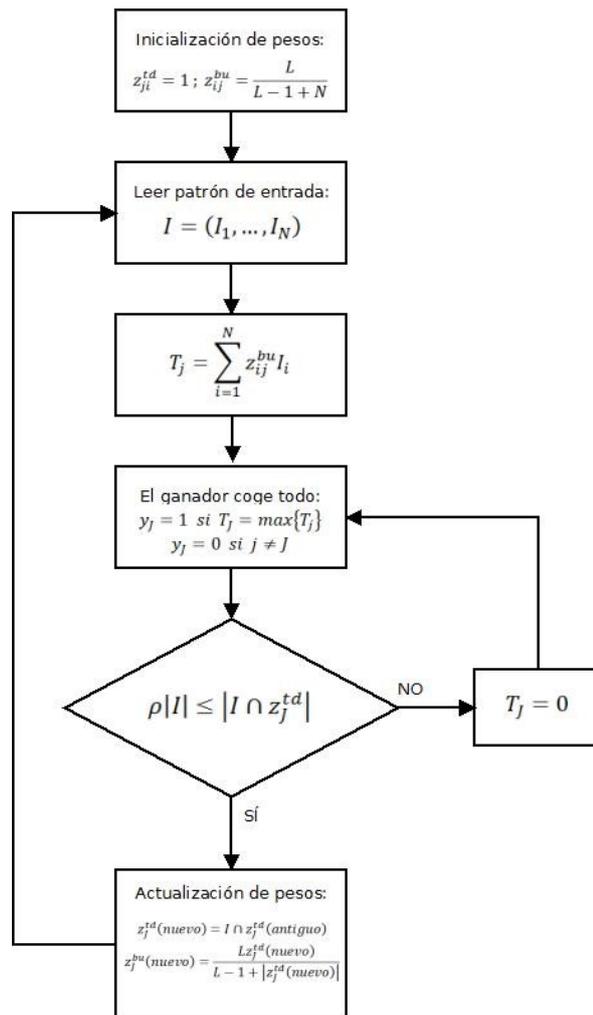
## B.1. ART1

ART1 se trata de una arquitectura de agrupamiento con entrenamiento no supervisado para patrones de entrada binarios. La arquitectura topológica se muestra en la Ilustración B.2, que aparece a continuación.



**Ilustración B.2:** Estructura topológica de la arquitectura ART1.

El siguiente diagrama de flujo se corresponde con el funcionamiento algorítmico de este tipo de arquitectura.



**Ilustración B.3:** Diagrama de flujo del funcionamiento algorítmico de ART1.

Sus principales propiedades son:

- Vigilancia grosera. Es la posibilidad de ajustar las categorías formadas de forma que se correspondan a una clasificación más o menos grosera, según convenga.
- Autoescalado. Esta propiedad permite a las características de un patrón en entrada ser tratadas como ruido o como señal, según el contexto. Si una característica, en la que difieren dos patrones, permite clasificar los patrones en clases distintas, se dice que esta característica es crítica, en caso contrario se dice que es ruido. El parámetro de vigilancia juega un papel fundamental para decidir si una característica es ruido o no.
- Autoestabilización rápida. Todos los pesos del sistema sujetos a un aprendizaje alcanzan un valor estacionario después de un número finito de presentaciones arbitrarias de patrones de entrada.
- Entrenamiento en línea. Para muchos algoritmos de agrupamiento, dado un conjunto de entrenamiento el algoritmo halla un número de categorías. Si se añade un nuevo patrón hace falta reiniciar todo el sistema y volver a entrenar el mismo, lo que se conoce como entrenamiento fuera de línea. En ART1 esto no pasa. Cada vez que se añade un nuevo patrón se contesta con una categoría comprometida o no y se actualiza esta categoría para incorporarla al conocimiento del sistema. Esta propiedad hace que los sistemas ART1 sean apropiados para realizar agrupaciones en tiempo real.
- Captura de eventos extraños. La ART1 puede aprender y formar agrupaciones de patrones de entrada que aparecen muy raramente. Gracias a la propiedad anterior el ART1 puede aprender un patrón extraño con tan solo un par de presentaciones del mismo. Un patrón raro será asociado a un nodo no comprometido, el cual no será alterado, ya que el resto de patrones difiere de este patrón extraño lo suficiente como para que esta categoría nunca resulte elegida.
- Acceso directo a familias de patrones de entrada. Al igual que el sistema cognitivo humano, ART1 tiene la habilidad de reconocer rápidamente un objeto que le es familiar al sistema, sin importarle el número de categorías que el sistema pueda tener almacenadas.
- Acceso directo a conjuntos y subconjuntos de patrones. Suponemos que tras un proceso de aprendizaje ART1 ha producido un conjunto de categorías. Supongamos también que de estas categorías hay dos,  $j_1$  y  $j_2$ , tales que  $z_{j_1}^{td} \subset z_{j_2}^{td}$ . En este caso se dice que  $j_1$  es una plantilla subconjunto de  $j_2$ , o equivalentemente,

$j_2$  es una plantilla superconjunto de  $j_1$ . Consideramos los siguientes dos patrones de entrada:

$$I^{(1)} = z_{j_1}^{td} \equiv \{z_{j_1 1}^{td}, z_{j_1 2}^{td}, \dots, z_{j_1 N}^{td}\} \quad (\text{Ecuación B.2})$$

$$I^{(2)} = z_{j_2}^{td} \equiv \{z_{j_2 1}^{td}, z_{j_2 2}^{td}, \dots, z_{j_2 N}^{td}\} \quad (\text{Ecuación B.3})$$

Esta propiedad asume que  $I^{(1)}$  tendrá acceso directo a  $j_1$  e  $I^{(2)}$  tendrá acceso directo a  $j_2$ .

- Predeterminación de la red a formar nuevas categorías. Sea cual sea el parámetro de vigilancia, el parámetro  $L$  predetermina la tendencia de formar un pequeño o gran número de categorías.

### B.2. ARTMAP

ARTMAP es una arquitectura de clasificación con entrenamiento supervisado para patrones de entrada binarios o parejas de clases de patrones. La topología de la arquitectura ARTMAP general puede verse en la Ilustración B.4.

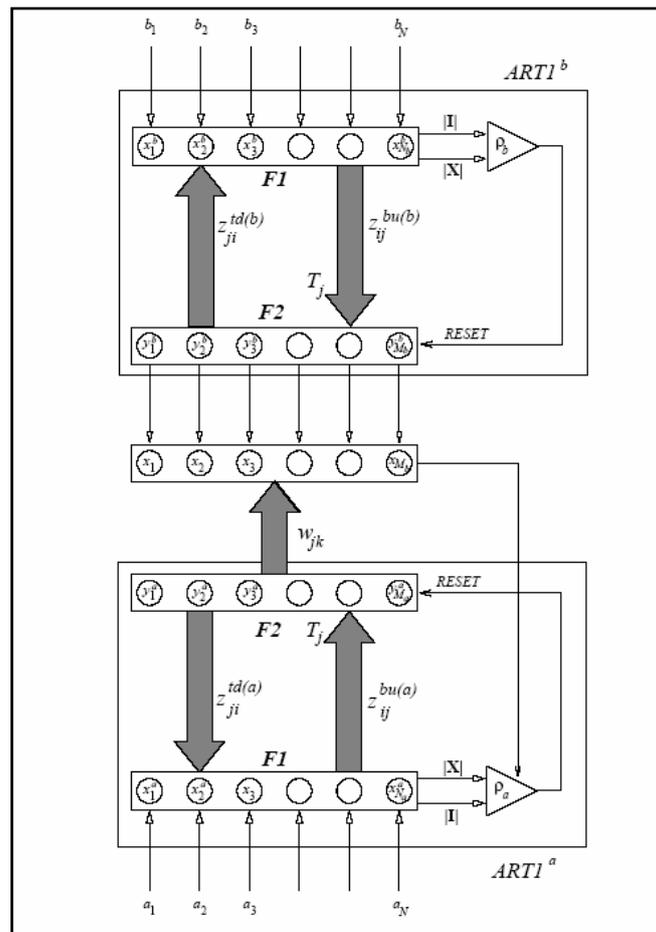


Ilustración B.4: Arquitectura ARTMAP general.

En el diagrama de flujo que se muestra a continuación podemos observar el funcionamiento algorítmico y su relación con la arquitectura ART1.

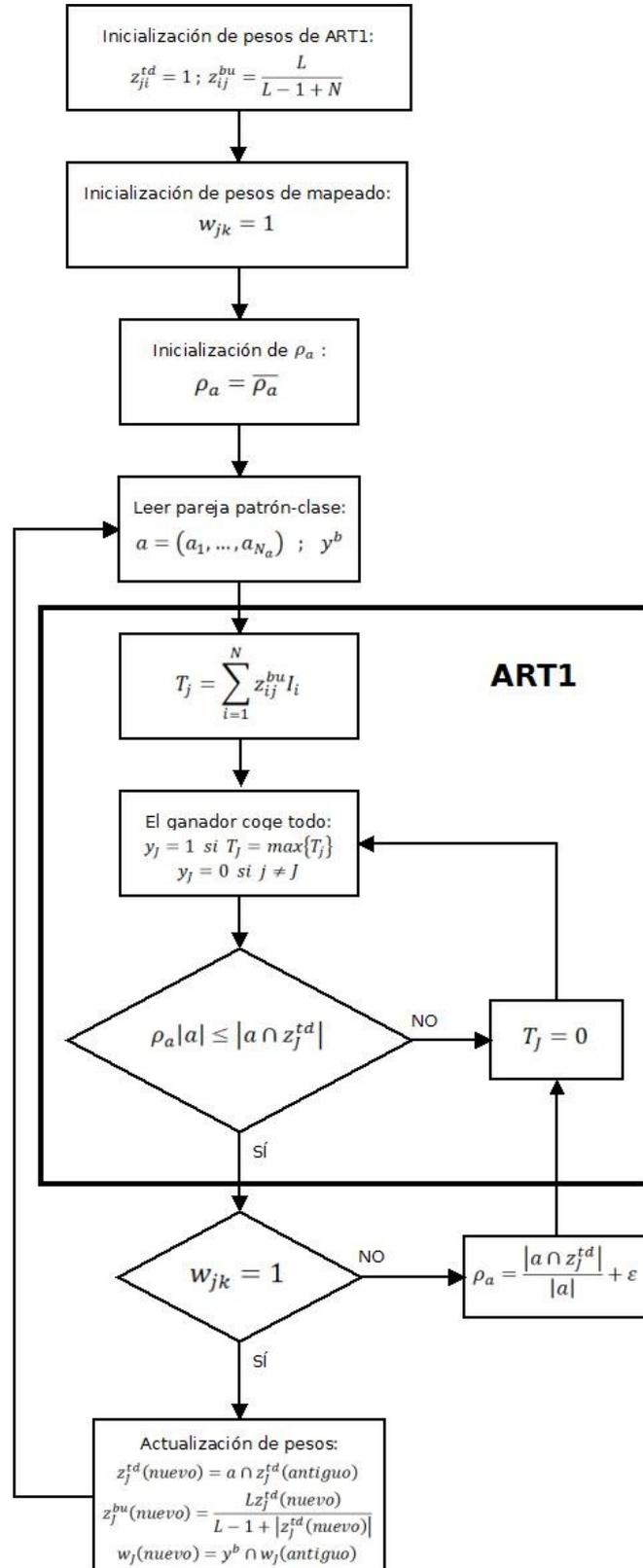


Ilustración B.5: Diagrama de flujo del funcionamiento algorítmico de la arquitectura ARTMAP.

ARTMAP hereda todas las propiedades de ART1, y además incluye las siguientes:

- Mayor capacidad de aprender eventos extraños. Es capaz de aprender eventos raros aunque estos tengan un mayor parecido a otros eventos más frecuentes, ya que si dos patrones muy similares tienen que ser englobados en clases distintas el subsistema *match-tracking* se encarga de aumentar el parámetro de vigilancia y conseguir la separación.
- Aprendizaje en pocos intentos. Aunque el entrenamiento no se haya estabilizado completamente, ARTMAP puede conseguir unas predicciones aceptables.
- Aprendizaje estable. Permite la actualización continua del conocimiento del sistema consiguiendo que se puedan aprender nuevas bases de datos sin eliminar el conocimiento previo.
- Incremento de la capacidad de memoria. Se puede aumentar la memoria sin sacrificar la estabilidad del aprendizaje rápido o la precisión de la capacidad de generalización.
- El sistema *match-tracking* tiene la habilidad de conjuntar máxima generalización y mínimo error de predicción en una prueba a prueba básica, usando solo operaciones locales.

### B.3. Fuzzy ART

Fuzzy ART es una arquitectura de agrupamiento con entrenamiento no supervisado para patrones de entrada analógicos. Como podemos observar en la siguiente figura, esta arquitectura es similar a la implementada en los sistemas ART1.

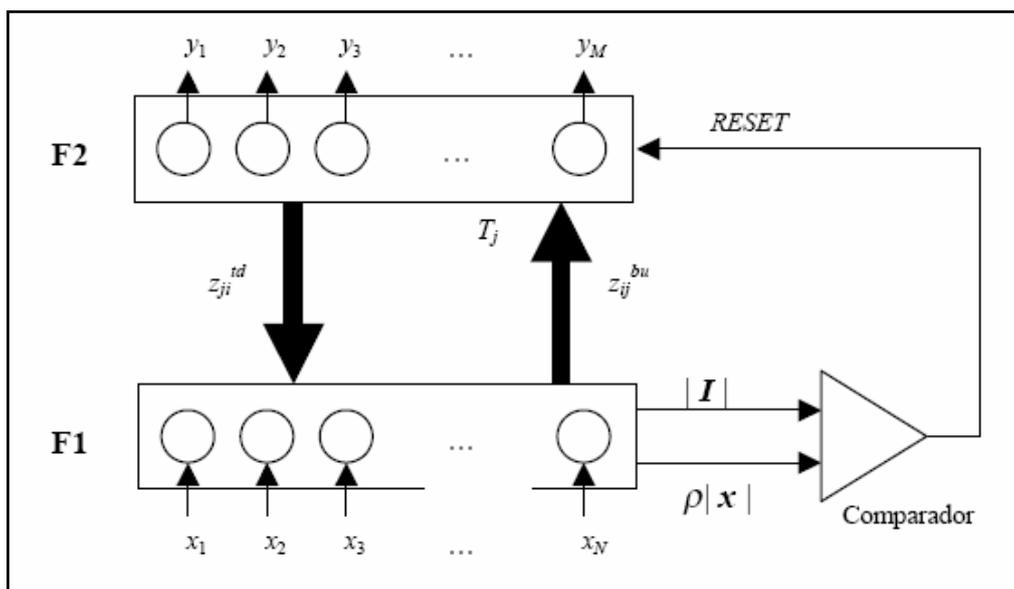
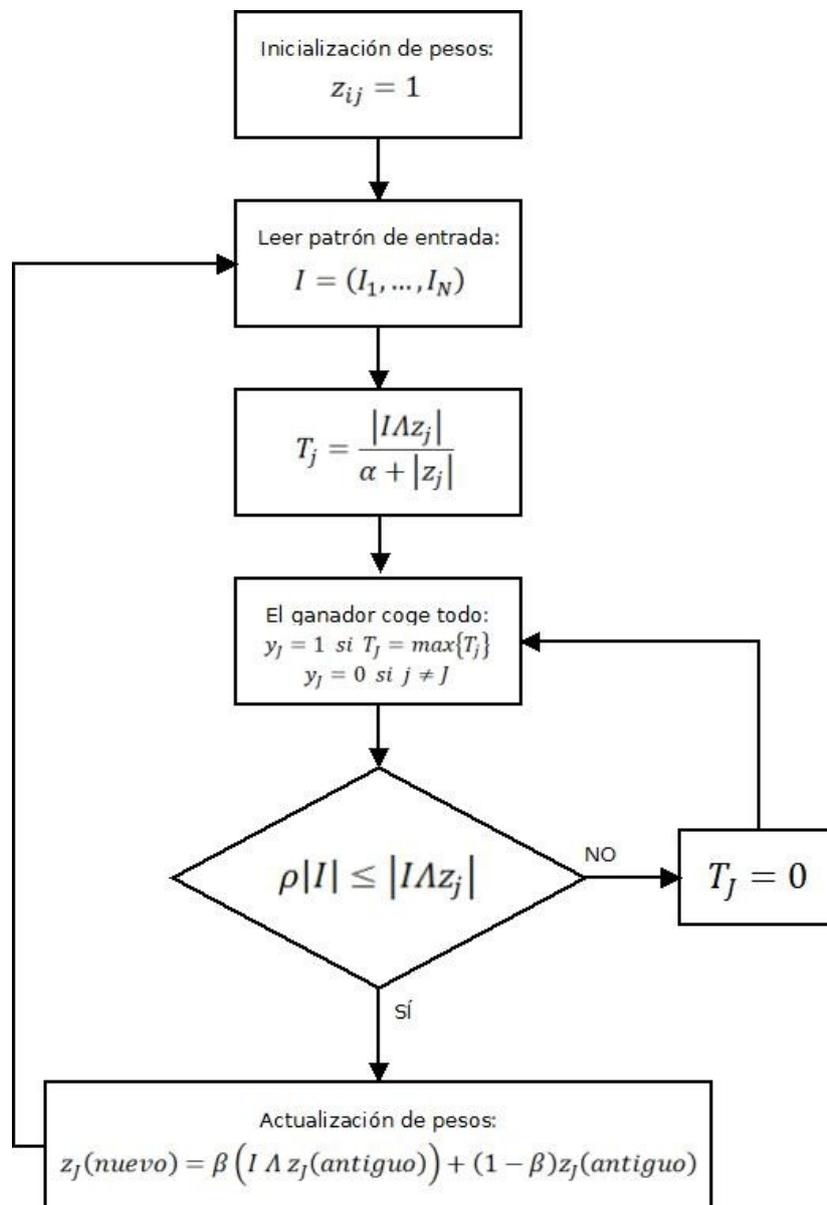


Ilustración B.6: Arquitectura funcional de un sistema Fuzzy ART.

Las principales diferencias entre la arquitectura ART1 y *Fuzzy* ART son las siguientes:

- Los patrones de entrada son analógicos. Esto quiere decir que  $I = (I_1, \dots, I_N)$  es un vector de dimensión  $N$  donde cada una de sus componentes  $I_i \in [0,1]$ .
- Solo existe un conjunto de pesos analógicos  $z_j = (z_{1j}, \dots, z_{Nj}) ; j = 1, \dots, M$ .
- En el cálculo de la función de elección  $T_j$  la operación de intersección  $\cap$  (AND binario) es sustituida por el operador confuso  $MIN \wedge$  (AND analógico).

En la Ilustración B.7 podemos ver el diagrama de flujo del algoritmo que sigue esta arquitectura.



**Ilustración B.7:** Diagrama de flujo del algoritmo utilizado por un sistema *Fuzzy* ART.

### B.4. Fuzzy ARTMAP

La red neuronal *Fuzzy ARTMAP* se trata de una arquitectura de clasificación con entrenamiento supervisado para patrones de entrada analógicos. Esta arquitectura es idéntica a la que se implementa en ARTMAP, excepto que se sustituye el módulo ART por una estructura *Fuzzy ART*. Consecuentemente, la única diferencia es que los patrones de entrada tienen carácter analógico. Estas categorías se representan por  $y^a$  e  $y^b$ , los cuales tienen todas sus componentes a 0 excepto una que es 1. Esto es igual que para la arquitectura ARTMAP, y también el módulo de mapa de campo se comporta de la misma manera en ARTMAP que en *Fuzzy ARTMAP*.

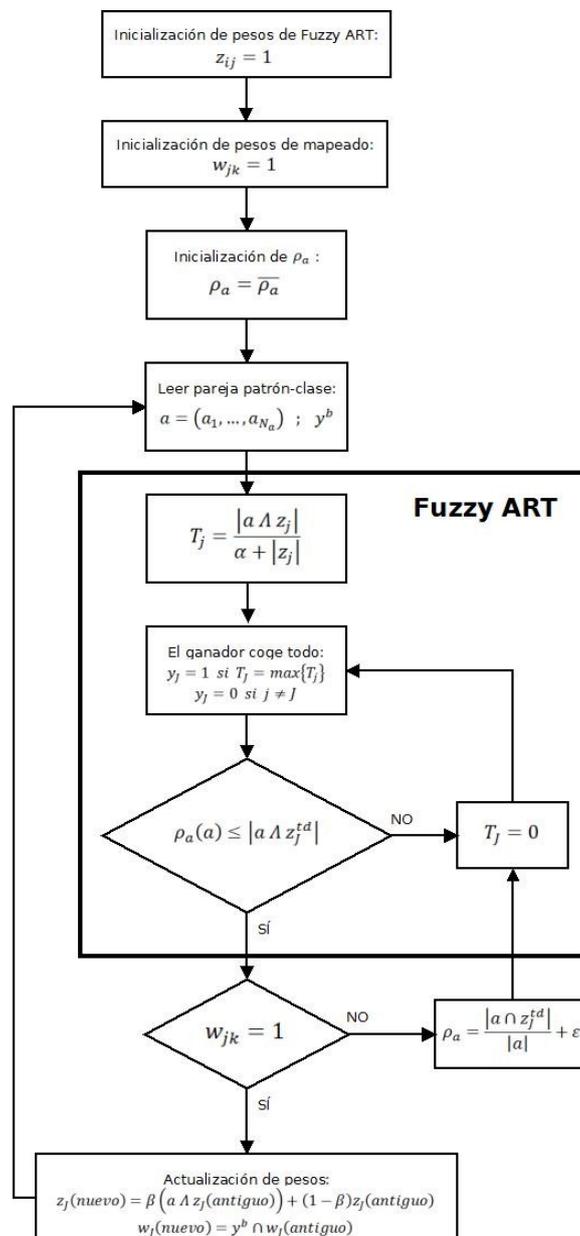


Ilustración B.8: Diagrama de flujo de la arquitectura *Fuzzy ARTMAP*.

Esta arquitectura incorpora dos modos de aprendizaje, el aprendizaje lento y el aprendizaje rápido, incluyendo un nuevo parámetro de vigilancia en el caso del aprendizaje lento que da mayor robustez al sistema. En cuanto a la estrategia de votación, *Fuzzy ARTMAP* consigue incrementar la precisión y obtiene unos resultados independientes del orden de presentación de los patrones de entrada. Esta arquitectura presenta las propiedades heredadas de las arquitecturas *Fuzzy ART* y *ARTMAP*, de las que destacamos las siguientes:

- Aprendizaje estable. Todos los pesos solo pueden decrecer a lo largo del proceso de entrenamiento.
- Solo son necesarios un pequeño número de parámetros, 2 en el caso del aprendizaje rápido y 5 en el caso de aprendizaje lento. El sistema siempre funcionará independientemente del valor de estos parámetros, siempre que estén dentro de sus rangos. Valores diferentes producen comportamientos diferentes.
- La arquitectura es la misma y los valores iniciales también son los mismos, independientemente del problema a resolver.
- Se crea el mínimo de categorías *Fuzzy-ART* necesarias para conseguir el criterio de precisión.
- Un valor pequeño de  $\bar{\rho}_\alpha$  consigue una mayor compresión de códigos y una mayor capacidad de generalización. Mayor  $\bar{\rho}_\alpha$  implica que cada categoría tendrá un menor tamaño y se esperará un menor error, aunque se generará un mayor número de categorías. Para  $\bar{\rho}_\alpha = 1$  cada patrón de entrada establecerá su propia categoría, lo que hace que *Fuzzy-ARTMAP* trabaje como un clasificador de mínima distancia.
- Como la red *ARTMAP*, *Fuzzy-ARTMAP* puede trabajar tanto fuera de línea como en línea.