Aprendizaje incremental de Sistemas Difusos Lingüísticos para Flujos de Datos

Javier Martín-Moreno[1], Francisco A. Márquez[1], Ana M. Roldán[3] y Antonio Peregrín[1,2]

[1] Centro de Estudios Avanzados en Física, Matemática y Computación (CEAFMC)
[2] Instituto Andaluz Interuniversitario en Ciencia de Datos e Inteligencia Computacional (DaSCI)
[3] Departamento de Tecnologías de la Información
Universidad de Huelva, Huelva, España
javier.martin@dti.uhu.es

Resumen. Este trabajo presenta una estrategia de aprendizaje en línea para Sistemas Difusos Lingüísticos Mamdani orientada a la predicción sobre flujos de datos continuos en contextos donde la interpretabilidad es prioritaria. El sistema propuesto adopta un enfoque de modelado difuso descriptivo, fundamentado en una única partición difusa global que garantiza la coherencia semántica de toda la Base de Reglas, junto con un mecanismo incremental de generación de reglas basado en cubrimiento. Para favorecer la cooperación entre reglas y limitar el crecimiento estructural, se incorpora un esquema de ponderación que regula su influencia y actúa como poda implícita. El ajuste en línea de etiquetas y pesos se realiza mediante un algoritmo genético sobre un búfer de aprendizaje reducido. Se evalúan tres estrategias de actualización del búfer: *First-In First-Out* (FIFO), *Reservoir Sampling* (RS), y *Sliding Window with Weight Decay* (SWWD). Esta propuesta se valida sobre dos conjuntos de datos de referencia, mostrando que el modelo de aprendizaje incremental en línea funciona, y que las políticas de búfer tienen efecto. A pesar de priorizar la interpretabilidad mediante una arquitectura Mamdani descriptiva, el método propuesto alcanza una precisión competitiva, constituyendo una alternativa relevante frente a modelos evolutivos más opacos.

Palabras Clave: Sistemas Difusos Evolutivos, Aprendizaje Evolutivo en Línea, Sistema Difuso Lingüístico, Aprendizaje Incremental.

1 Introducción

El análisis de flujos de datos en tiempo real es clave en aplicaciones modernas que requieren monitorización continua, como las infraestructuras del Internet de las Cosas (IoT) y las plataformas de gestión energética. La característica principal de este tipo de datos es su naturaleza no estacionaria y temporalmente cambiante, manifestada como *concept drift*, es decir, variaciones en las propiedades estadísticas de la función objetivo a lo largo del tiempo. Los modelos predictivos implementados en estos contextos deben ser capaces de adaptarse de manera continua en tiempo real; de lo contrario, su rendimiento puede degradarse. Además, el elevado volumen y la velocidad de la generación de datos exigen estrategias de aprendizaje en línea eficientes de baja latencia.

Si bien los avances recientes en aprendizaje automático, especialmente los basados en arquitecturas profundas, han logrado una capacidad predictiva sobresaliente, lo hacen habitualmente a expensas de la explicabilidad. En numerosas aplicaciones reales, particularmente aquellas en las que la seguridad, la confianza y la transparencia son requisitos fundamentales, resulta imprescindible disponer de modelos que preserven su interpretabilidad y sean comprensibles para los expertos a lo largo de su ciclo de vida.

En este contexto, los Sistemas Difusos que Evolucionan (*Evolving Fuzzy System* - EFS) [1] ofrecen un equilibrio prometedor: combinan capacidades de aprendizaje incremental con representaciones lingüísticas inherentemente interpretables. Los modelos EFS clásicos aprenden y ajustan dinámicamente reglas difusas a medida que llegan nuevos datos, usando mecanismos de generación, fusión y eliminación de reglas con el objetivo de mantener la compacidad y la relevancia de la Base de Conocimiento (BC). Sin embargo, a pesar de ser más interpretables que los modelos de caja negra, muchos EFS se basan en arquitecturas de tipo Takagi-Sugeno (TS) o Neuro-Difusos (ND), donde los consecuentes son habitualmente funciones lineales y los conjuntos difusos difieren entre reglas. Aunque este tipo de diseño ofrece una mayor capacidad predictiva, reduce la interpretabilidad semántica, considerándose modelos de caja gris.

En el paradigma actual de la Inteligencia Artificial Explicable (*Explainable Artificial Intelligence* - XAI) [2], la interpretabilidad se concibe como una propiedad fundamental de los sistemas desplegados en entornos de alta sensibilidad. Así, las arquitecturas difusas lingüísticas de tipo Mamdani [3] resultan especialmente atractivas, dado que el empleo de etiquetas lingüísticas en los consecuentes produce estructuras de reglas significativamente más accesibles a la interpretación humana. En particular, los modelos descriptivos, en los que todas las reglas comparten una única partición difusa global en la Base de Datos (BD), garantizan la consistencia semántica en toda la Base de Reglas (BR), propiedad ausente en las alternativas aproximativas, propias de los EFS.

No obstante, adoptar una arquitectura Mamdani en un entorno dinámico plantea dos retos centrales: el crecimiento de la BR, puesto que la creación incremental de reglas puede generar fácilmente redundancias y solapamientos; y la cooperación entre reglas, ya que las nuevas reglas incorporadas pueden interferir de forma no deseada con las previamente existentes, deteriorando el rendimiento predictivo global [4].

Para abordar estos desafíos, se propone un sistema difuso Mamdani descriptivo que combina un método incremental de generación de reglas por cubrimiento, un mecanismo de ponderación de reglas que favorece su cooperación y proporciona una poda implícita, y una técnica de ajuste en línea de las etiquetas difusas y de los pesos de las reglas mediante un algoritmo evolutivo que opera sobre un búfer de memoria limitado.

El objetivo de este trabajo no es superar la precisión de los EFS de tipo TS y ND, sino ofrecer una alternativa transparente y lingüísticamente coherente con capacidad de adaptación en línea. Para fortalecer el mecanismo de adaptación, se evalúan tres estrategias de actualización del búfer, analizando el impacto de cada política sobre el rendimiento del modelo en entornos dinámicos, en respuesta a una cuestión abierta [5].

El resto del artículo se organiza de la siguiente manera: la Sección II revisa los antecedentes más relevantes. La Sección III formula el problema. La Sección IV describe el modelo propuesto. La Sección V detalla el estudio experimental. La Sección VI discute los resultados obtenidos. La Sección VII recoge las conclusiones del trabajo.

2 Antecedentes

Esta sección revisa las contribuciones más relevantes del ámbito EFS y sitúa la propuesta presentada dentro del marco científico.

Desde el trabajo seminal de Angelov [6] a principios de la década de los 2000, los EFS han sido reconocidos por su capacidad de adaptar la estructura de los modelos difusos en línea mediante aprendizaje en una sola pasada, considerándose adecuados para aplicaciones en tiempo real. Gran parte de los modelos EFS propuestos hacen uso de arquitecturas TS [7]-[12] o ND [13]-[16], compartiendo dos características comunes:

- Generación de reglas mediante agrupamiento basado en densidad, donde los conjuntos difusos son habitualmente de tipo gaussiano independientes para cada regla.
- Adaptación de los consecuentes por Mínimos Cuadrados Recursivos (*Recursive Least Squares* - RLS), primando la precisión en detrimento de la interpretabilidad.

La definición independiente de los conjuntos difusos de los antecedentes en cada regla genera particiones semánticas heterogéneas, de modo que cada regla gestiona términos lingüísticos distintos. Si bien esta flexibilidad mejora la precisión, debilita la interpretabilidad y la transparencia del modelo resultante. Asimismo, el ajuste de los parámetros de los consecuentes mediante variantes del algoritmo RLS dificulta en gran medida la comprensión del proceso de aprendizaje interno, haciendo que el proceso de inferencia del modelo sea poco transparente y dificultando la explicabilidad a posteriori.

Por el contrario, las arquitecturas Mamdani emplean consecuentes lingüísticos y, cuando se combinan con una partición difusa global, preservan la consistencia semántica en toda la BR, simplificando tanto la explicación como la inspección humana. Estos modelos están típicamente asociados a BCs descriptivas, en las que todas las reglas de la BR comparten la misma partición difusa global definida en una BD común. Esta estructura compartida garantiza que los términos lingüísticos mantengan una coherencia semántica uniforme en toda la BR, reforzando así su interpretabilidad global.

Dado que estos modelos presentan un menor número de parámetros ajustables y una estructura con menor grado de libertad, habitualmente se emplean técnicas de optimización evolutiva para su ajuste. Los algoritmos genéticos (AG) han sido ampliamente aplicados en este contexto, gracias a su capacidad de explorar espacios de búsqueda complejos y no lineales sin necesidad de información sobre el gradiente, proporcionando un buen equilibrio entre exploración del espacio de búsqueda y explotación de regiones prometedoras, ofreciendo buenos resultados con un alto coste computacional.

Sin embargo, el ajuste evolutivo se ha llevado a cabo tradicionalmente de forma *offline*, sobre conjuntos de datos estáticos mediante procesamiento iterativo. En consecuencia, muchos sistemas difusos de tipo Mamdani carecen de la capacidad de adaptarse a los cambios del entorno en tiempo real. En entornos dinámicos es imprescindible el uso de estrategias de ajuste en línea, siendo éste el principal desafío: los algoritmos evolutivos requieren múltiples evaluaciones de la función de evaluación, lo que generalmente implica la necesidad de más de una observación en cada instante de tiempo. Esto motiva el uso de un búfer de aprendizaje, cuya gestión se convierte en un aspecto clave de la adaptación en línea.

3 Formulación del Problema

El objetivo es desarrollar un método capaz de aprender y ajustar en tiempo real modelos de predicción sobre flujos de datos, donde la información llega de forma secuencial y puede sufrir *concept drift*.

En numerosos escenarios de regresión y predicción de series temporales, los modelos de aprendizaje se apoyan en técnicas de memoria explícita, de modo que una secuencia de observaciones pasadas se utiliza como entrada en lugar de depender exclusivamente de la observación actual. Siendo x_i el elemento de entrada en el instante t_i, resulta necesario mantener un pequeño búfer de retardos $\hat{x}_i$ de tamaño λ que almacena las observaciones más recientes. El vector de entrada se define como:

$$\hat{x}_i = [x_{i-\lambda}, \dots, x_i]$$

y el modelo trata de predecir el siguiente valor, típicamente:

$$y_i \approx x_{i+1}$$

Esta configuración define un sistema de memoria explícita, en el que la información pasada se proporciona directamente al modelo, permitiéndole capturar las dependencias temporales a corto plazo esenciales para una predicción precisa en entornos dinámicos. Por otro lado, los sistemas de memoria implícita almacenan la información en estados o parámetros internos y pueden realizar predicciones empleando únicamente la observación actual.

El rendimiento de los modelos predictivos empleados en tareas de regresión y predicción de series temporales se evalúa mediante dos métricas fundamentales. El Error Cuadrático Medio (*Root Mean Squared Error*, RMSE), definido en la ecuación (1), cuantifica el error absoluto de predicción en las mismas unidades que la variable objetivo y. Por su parte, el Índice de Error No Dimensional (*Non-Dimensional Error Index*, NDEI), dado en la ecuación (2), proporciona una medida de error normalizada y adimensional, que facilita la comparación justa del rendimiento de los modelos entre conjuntos de datos con diferentes escalas y varianzas.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(f(\hat{x}_i) - y_i)^2} \tag{1}$$

donde N es el número de instancias evaluadas, $f(\cdot)$ representa la predicción inferida por el modelo, $\hat{x}_i$ es el valor del búfer de retardos en el instante i, e y_i es el valor objetivo (típicamente x_{i+1}).

$$NDEI = \frac{RMSE}{\sigma(\hat{x})} \tag{2}$$

donde $\sigma(\hat{x})$ es la desviación típica del conjunto de búferes de retardos evaluado.

4 Modelo Propuesto

Esta sección ofrece una descripción detallada de los principios de diseño y los aspectos de implementación que definen la estructura y el funcionamiento del modelo propuesto.

4.1 Arquitectura del Modelo

El sistema propuesto adopta un Sistema Basado en Reglas Difusas (SBRD) de tipo Mamdani descriptivo, compuesto por:

- Una BD difusa global: cada variable de entrada y de salida se particiona mediante funciones de pertenencia triangulares que cubren el dominio conocido de la variable. Estas funciones de pertenencia se corresponden con conceptos lingüísticos fijos (p. ej., Bajo, Medio, Alto) comunes a todas las reglas. Las posiciones y anchuras de estas etiquetas se actualizan en línea a medida que evoluciona la distribución de los datos, partiendo de una distribución equitativa de todo el dominio.
- Una BR lingüística: las reglas emplean etiquetas lingüísticas tanto en antecedentes como en consecuentes, siguiendo la estructura canónica Mamdani:
  $$SI\ x_1\ es\ \mu_a\ Y\ \ldots Y\ x_n\ es\ \mu_b\ ENTONCES\ y'\ es\ \mu_c$$
  Dado que todas las reglas hacen referencia a la misma BD compartida, la BC resultante mantiene su coherencia semántica.
- Un mecanismo de ponderación de reglas: cada regla incorpora un peso (*Rule Weight* - RW $\in$ [0,1]) que modula su contribución durante la inferencia. Las reglas con poco peso se vuelven inactivas en la práctica, funcionando como poda implícita para controlar la redundancia y reforzar la cooperación entre las reglas con mayor peso, mejorando la precisión global del modelo.

Tradicionalmente, este tipo de BC lingüística se combina con un mecanismo de inferencia que emplea fuzzificación puntual, la T-norma mínimo para el emparejamiento e inferencia de reglas, y la estrategia FITA (*First Infer, Then Aggregate*) con defuzzificación mediante la media ponderada de máximos.

Por regla general, la generación de la BR por cubrimiento y el ajuste evolutivo de la BD se llevan a cabo mediante técnicas de procesamiento por lotes, iterando sobre todos los datos de entrenamiento disponibles. Sin embargo, el enfoque propuesto en este trabajo busca adaptar dichas técnicas a entornos dinámicos y continuamente cambiantes, donde tanto la generación de reglas como el ajuste de parámetros deben ejecutarse en línea, sin depender de datos estáticos obtenidos de forma *offline*.

4.2 Aprendizaje y Adaptación en Línea

A medida que los datos llegan de forma secuencial, el SBRD se actualiza de forma continua:

1. Generación de reglas por cubrimiento: Si el nuevo elemento observado no está suficientemente cubierto por ninguna regla existente, se genera una nueva regla utilizando la BD compartida.
2. Ajuste evolutivo en línea: El AG elegido, de tipo CHC, se emplea para ajustar la distribución de las etiquetas lingüísticas y los pesos de las reglas. Otros elementos son:
   a. Codificación: el cromosoma usa la codificación 3-tuplas (posición, anchura, etiqueta) para las etiquetas de la BD codificadas como valores reales, y los pesos de las reglas codificados como reales en el intervalo [0,1], inicializados a 1.
   b. Selección elitista para preservar los mejores cromosomas.
   c. Prevención de incesto mediante umbral de distancia de Hamming sobre una representación de punto fijo, para evitar la convergencia prematura.
   d. Operador de cruce BLX-α.
   e. Técnica de reinicio de la población para prevenir el estancamiento.
   f. Función de evaluación NDEI.

De este modo, el algoritmo optimiza la redistribución de las etiquetas lingüísticas teniendo en cuenta su influencia colectiva sobre todas las reglas, favoreciendo así la cooperación entre ellas, un aspecto clave para alcanzar una mayor precisión del modelo.

Los AGs necesitan un conjunto de muestras de tamaño considerable para evaluar los cromosomas. En entornos dinámicos, no es eficaz iterar sobre una única observación de forma significativa. En consecuencia, no resulta factible llevar a cabo el ajuste evolutivo en tiempo real bajo un enfoque estricto de pasada única. Para solucionar esta limitación, se mantiene un búfer de aprendizaje de tamaño limitado que conserva un historial reducido de datos recientes.

Este búfer permite ejecutar el AG en línea cada vez que llega una nueva observación, no evaluándola de forma aislada, sino apoyándose en un conjunto compacto de elementos almacenados previamente de manera temporal. El búfer contiene así un número reducido de muestras pasadas sobre las que los cromosomas pueden evaluarse de forma iterativa. Para poblar y mantener este búfer, se exploran tres políticas de actualización:

- *First In, First Out* (FIFO): Al recibir un nuevo elemento, se descarta el más antiguo del búfer. Conserva únicamente las observaciones más recientes. Rápida pero potencialmente limitada ante ciertas variaciones en los patrones.
- *Reservoir Sampling* (RS): Este enfoque preserva una muestra estadísticamente imparcial en el tiempo de todos los datos pasados. Cada vez que llega un nuevo elemento, cada elemento existente tiene la misma probabilidad uniforme de ser reemplazado. Ofrece una mayor robustez frente a cambios a largo plazo y cambios de patrones recurrentes.
- *Sliding Window with Weight Decay* (SWWD): Se asigna un peso decreciente a las instancias más antiguas, de modo que los elementos con mayor antigüedad tienen mayor probabilidad de ser reemplazados. Además, este peso influye en la evaluación de los cromosomas, reduciendo la contribución de las instancias más antiguas en el proceso de validación. Como resultado, las observaciones de mayor antigüedad se vuelven progresivamente más susceptibles de ser reemplazadas y menos influyentes en el proceso de ajuste.

La Figura 1 ilustra el flujo general de la metodología de ajuste en línea propuesta. Para cada nuevo búfer de retardos de entrada $\hat{x}_{i-2}$, el búfer de aprendizaje se actualiza de acuerdo con la política seleccionada. Dicho búfer sustenta a continuación el ajuste incremental de la BC del SBRD: se generan nuevas reglas Mamdani según sea necesario y, posteriormente, los términos lingüísticos y los pesos de las reglas se ajustan evaluando cada cromosoma sobre el contenido actual del búfer de aprendizaje.

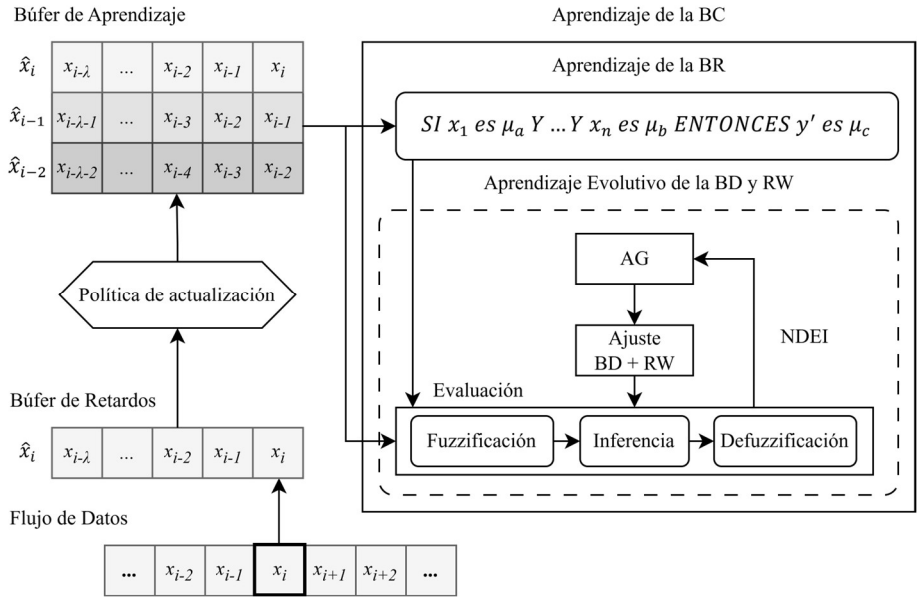


Fig. 1. Flujo del proceso de ajuste en línea del sistema difuso lingüístico.

5 Estudio Experimental

En esta sección se presenta el estudio experimental llevado a cabo para evaluar el rendimiento del modelo propuesto.

5.1 Configuración Experimental

Para evaluar la precisión del modelo se han seleccionado dos conjuntos de datos de referencia de uso extendido, considerados representativos en la literatura sobre EFS.

- Mackey-Glass: este conjunto de datos corresponde a una serie temporal no lineal generada a partir de la ecuación diferencial con retardo mostrada en (3).

$$\frac{dx\,(t)}{dt} = \frac{0.2x(t-\tau)}{1+x^{10}(t-\tau)} - 0.1x(t) \tag{3}$$

Se emplearon los parámetros habituales de esta serie: condición inicial $x(0) = 1.2$ y parámetro de retardo $\tau = 17$, generando un total de 6.000 muestras. El búfer de retardos se configuró con el vector de entrada $[x_{(t-18)}, x_{(t-12)}, x_{(t-6)}, x_{(t)}]$, para predecir el valor futuro $y = x_{(t+85)}$. Desde $t = 201$ a $t = 3.200$ se usaron para entrenamiento, y desde $t = 5.001$ a $t = 5.500$ para test.

- S&P500: este conjunto de datos corresponde a la serie temporal definida por los precios de cierre diarios del índice S&P500, obtenidos del repositorio de *Yahoo! Finance*, cubriendo el periodo desde el 3 de Enero de 1950 al 12 de Marzo de 2009, compuesto por 14.893 instancias, concatenado con su versión homóloga inversa en el tiempo, generando un conjunto de datos con 29.786 observaciones en total. El búfer de retardos se configuró con el vector de entrada $[x_{(t-4)}, x_{(t-3)}, x_{(t-2)}, x_{(t-1)}, x_{(t)}]$ para predecir el valor futuro $y = x_{(t+1)}$. Desde $t = 0$ a $t = 14.892$ se usaron para entrenamiento, y desde $t = 14.893$ a $t = 29.892$ para test.

Los hiperparámetros del sistema difuso evolutivo propuesto han sido los siguientes: todas las variables de entrada y salida se particionaron en tres etiquetas lingüísticas triangulares simétricas, definiendo la granularidad de la partición difusa global de la BD. Para el ajuste evolutivo, el tamaño de la población se fijó en 10 cromosomas en el modelo incremental propuesto, y 50 cromosomas en el modelo entrenado con todos los datos (modelo iterativo *offline*), que se utilizó como referencia para la comparación. El número máximo de evaluaciones por paso de aprendizaje evolutivo se limitó a 10 en la configuración en línea y 1.000 en el modelo *offline* de referencia. Se empleó el operador de cruce BLX-α con $\alpha = 0.5$. Para preservar la diversidad genética, se aplicó la prevención de incesto permitiendo el cruce únicamente entre cromosomas cuya distancia de Hamming superara ($Longitud \times 64 / 4$). Adicionalmente, se aplicó una estrategia de elitismo que conserva el 10% de los mejores cromosomas al reiniciar la población.

Con el fin de caracterizar el comportamiento del modelo propuesto, se diseña un estudio comparativo estructurado en dos etapas. A modo de ablación interna, se analiza en primer lugar el impacto de las políticas de actualización del búfer (FIFO, RS y SWWD) sobre el rendimiento predictivo. Se contrastan los resultados para distintos tamaños de búfer (100, 150, 300 y 600 instancias) con dos referencias de los extremos del espectro de aprendizaje incremental: el modelo sin búfer, entrenado en pasada única sobre cada observación sin posibilidad de reevaluación, y el modelo iterativo *offline*, entrenado con la totalidad de los datos de entrenamiento disponibles mediante procesamiento por lotes, que actúa como cota superior de referencia. En la segunda etapa se compara el mejor resultado del modelo propuesto frente a enfoques EFS del estado del arte reportados en la literatura.

5.2 Resultados

Para evaluar la influencia de las distintas políticas de búfer de aprendizaje, las Tablas 1 y 2 recogen los resultados obtenidos por el modelo difuso Mamdani descriptivo propuesto cuando es ajustado en línea mediante las diferentes políticas de búfer con diversos tamaños de búfer sobre la serie temporal Mackey-Glass y S&P500 respectivamente.

La Tabla 3 contextualiza los resultados frente a enfoques EFS del estado del arte.

Tabla 1. Comparativa de Políticas de Búfer, evaluadas mediante NDEI sobre Mackey-Glass

Tamaño de Búfer	Política de Búfer			Sin Política
	FIFO	RS	SWWD	
1 (pasada única)	-	-	-	1.324
100	1.040	1.469	1.212	-
150	1.224	1.215	1.125	-
300	1.614	0.824	1.205	-
600	1.451	0.698	0.836	-
3.200 (*offline*)	-	-	-	0.875

Tabla 2. Comparativa de Políticas de Búfer, evaluadas mediante NDEI sobre S&P500

Tamaño de Búfer	Política de Búfer			Sin Política
	FIFO	RS	SWWD	
1 (pasada única)	-	-	-	1.071
100	0.465	0.934	0.189	-
150	1.349	0.161	0.184	-
300	1.221	0.070	0.341	-
600	0.291	0.059	0.256	-
14.893 (*offline*)	-	-	-	0.067

Se observa que las políticas de búfer que retienen información a más largo plazo tienden a obtener mejores resultados, siempre que el tamaño del búfer sea suficientemente grande para que el AG pueda ajustar eficazmente los parámetros de la BC. En ocasiones, esta ventaja resulta lo suficientemente significativa como para que la configuración en línea supere la precisión obtenida mediante aprendizaje iterativo offline.

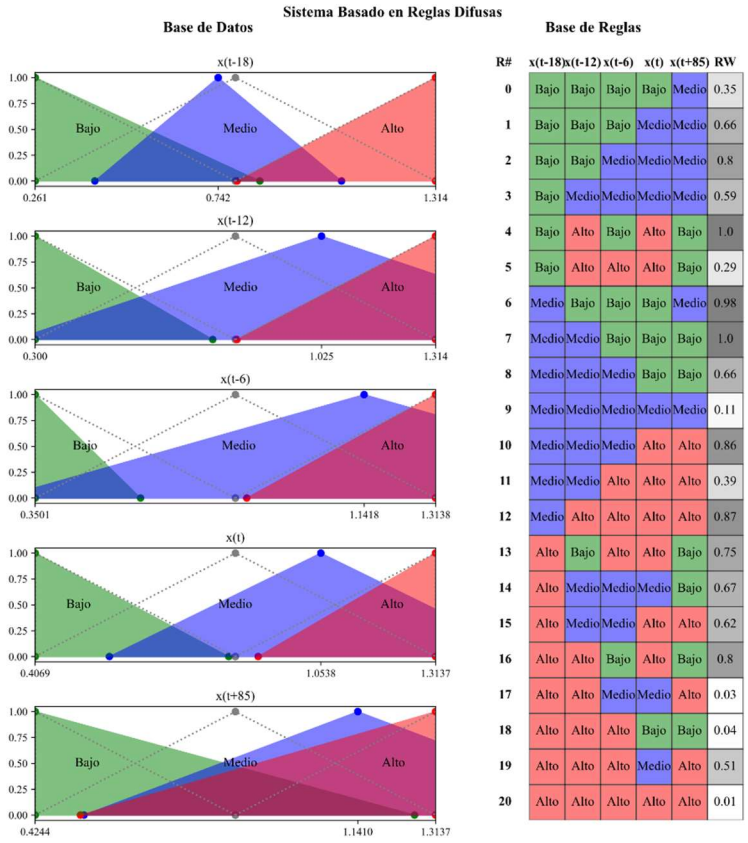
Tabla 3. Comparativa de precisión frente a EFS del Estado del Arte

Model	Mackey-Glass (NDEI)	S&P 500 (NDEI)
eTS [7]	0.347	0.015
Simpl_eTS [8]	0.279	0.045
LEOA [9]	0.248	0.123
ALMMo [10]	0.402	0.015
SEFS [11]	0.375	0.018
EFS-SLAT [12]	0.139	0.016
EFuNN [13]	0.139	0.154
DENFIS [14]	0.278	0.020
PANFIS [15]	0.301	0.090
GENEFIS [16]	0.280	0.070
RS (600)	0.698	0.059

A pesar de su diseño orientado a la interpretabilidad, el modelo propuesto alcanza resultados aceptables, especialmente teniendo en cuenta las limitaciones inherentes, como el empleo de una BD descriptiva en lugar de aproximativa y una BR de tipo Mamdani en lugar de TS o ND.

Cabe destacar que el objetivo de este trabajo no es maximizar la precisión predictiva, sino ofrecer una alternativa más interpretable que preserve la consistencia lingüística manteniendo al mismo tiempo un nivel de exactitud razonable.

La Figura 2 ilustra la BC obtenida utilizando la política RS con un tamaño de 600 sobre Mackey-Glass. Las etiquetas difusas iniciales se representan como funciones triangulares punteadas en gris, distribuidas simétricamente sobre el dominio completo de cada variable. Tras el proceso de ajuste, las etiquetas difusas adaptadas, mostradas en verde, azul y rojo, presentan posiciones y anchuras actualizadas, reflejando los cambios introducidos por el proceso evolutivo. Finalmente, los pesos de las reglas, inicializados todos a 1, han sido adaptados durante el aprendizaje, obteniendo nuevos valores que modulan la contribución de cada regla en la inferencia final.



Base de Reglas

R#	x(t-18)	x(t-12)	x(t-6)	x(t)	x(t+85)	RW
0	Bajo	Bajo	Bajo	Bajo	Medio	0.35
1	Bajo	Bajo	Bajo	Medio	Medio	0.66
2	Bajo	Bajo	Medio	Medio	Medio	0.8
3	Bajo	Medio	Medio	Medio	Medio	0.59
4	Bajo	Alto	Bajo	Alto	Bajo	1.0
5	Bajo	Alto	Alto	Alto	Bajo	0.29
6	Medio	Bajo	Bajo	Bajo	Medio	0.98
7	Medio	Medio	Bajo	Bajo	Bajo	1.0
8	Medio	Medio	Medio	Bajo	Bajo	0.66
9	Medio	Medio	Medio	Medio	Medio	0.11
10	Medio	Medio	Medio	Alto	Alto	0.86
11	Medio	Medio	Alto	Alto	Alto	0.39
12	Medio	Alto	Alto	Alto	Alto	0.87
13	Alto	Bajo	Alto	Alto	Bajo	0.75
14	Alto	Medio	Medio	Medio	Bajo	0.67
15	Alto	Medio	Medio	Alto	Alto	0.62
16	Alto	Alto	Bajo	Alto	Bajo	0.8
17	Alto	Alto	Medio	Medio	Alto	0.03
18	Alto	Alto	Alto	Bajo	Bajo	0.04
19	Alto	Alto	Alto	Medio	Alto	0.51
20	Alto	Alto	Alto	Alto	Alto	0.01

Fig. 2. BC resultante del modelo propuesto tras el ajuste de la BD y los pesos de las reglas.

Cabe destacar que el resultado permanece siempre auditable. A diferencia de los consecuentes de las arquitecturas TS o ND, la BC conserva en todo momento una representación lingüística legible, permitiendo inspeccionar cómo evolucionan los conceptos y las reglas a lo largo del tiempo. Esta trazabilidad permite la auditoría de los modelos.

6 Discusión

En términos generales, los resultados indican que las dos políticas de actualización del búfer de aprendizaje más avanzadas comienzan a mostrar su eficacia a partir de un determinado tamaño. Este comportamiento era previsible, dado que el búfer tiene sentido precisamente como mecanismo para preservar una memoria de experiencia (ejemplos históricos) que orienta el aprendizaje, incorporando lo más relevante de la dinámica pasada del sistema y consolidando progresivamente su BC.

Sin embargo, uno de los hallazgos más destacados es que el modelo propuesto, cuando se ajusta mediante una política de actualización sofisticada sobre un búfer de tamaño fijo, alcanza una precisión predictiva en ocasiones superior a la obtenida entrenando el mismo modelo sobre el conjunto de datos completo. Esto evidencia que mantener un búfer acotado pero representativo puede resultar más beneficioso que utilizar la totalidad del historial de datos, mejorando además la eficiencia del AG. Es decir, la representatividad del búfer importa más que la exhaustividad del registro histórico.

A la vista de estos resultados, conviene señalar algunas limitaciones del enfoque propuesto: la naturaleza descriptiva del modelo, junto con las restricciones inherentes a las reglas de tipo Mamdani, conduce inevitablemente a una precisión predictiva inferior a la de las arquitecturas TS o ND. Además, la optimización evolutiva es computacionalmente más costosa que los métodos basados en gradiente, un aspecto relevante que debe valorarse cuidadosamente en entornos de flujo de datos en tiempo real.

7 Conclusiones

En este trabajo se ha presentado un enfoque de aprendizaje en línea para sistemas difusos lingüísticos Mamdani de tipo descriptivo, que combina el aprendizaje incremental de reglas con el ajuste evolutivo de las particiones difusas globales y los pesos de las reglas, sobre un búfer de aprendizaje de tamaño limitado. Se han comparado tres políticas distintas de actualización del búfer (FIFO, RS y SWWD) con el fin de analizar su impacto en el aprendizaje evolutivo en línea.

El estudio experimental realizado sobre los conjuntos de datos de referencia pone de manifiesto que la composición y la dinámica de actualización del búfer de aprendizaje desempeñan un papel determinante en la capacidad de adaptación en línea del modelo. En particular, las políticas que preservan una representación más rica de la información pasada obtienen un mejor rendimiento, logrando una precisión competitiva con respecto a los modelos difusos evolutivos orientados a la exactitud, sin renunciar a la alta interpretabilidad inherente a los modelos Mamdani descriptivos. Este equilibrio adquiere una importancia creciente en despliegues sobre flujos de datos, donde los modelos deben adaptarse con el tiempo y mantenerse auditables y comprensibles.

Como líneas de trabajo futuro, se contempla explorar un abanico más amplio de estrategias de gestión del búfer, evaluar la aplicabilidad del modelo propuesto en otros escenarios reales de flujo de datos y añadir métricas de error que evalúen mejor el ajuste del modelo.

Agradecimientos. Este trabajo ha sido financiado por el Ministerio de Ciencia, Innovación y Universidades, bajo el proyecto PID2023-150070NB-I00.

Declaración de Intereses. Los autores no tienen conflicto de interés que declarar.

Referencias

1. Angelov, P.: Evolving fuzzy systems. In: Granular, Fuzzy, and Soft Computing, pp. 725–741. Springer US, New York (2023)
2. Herrera, F.: Reflections and attentiveness on explainable artificial intelligence (XAI): The journey ahead from criticisms to human-AI collaboration. Inf. Fusion 121, 103133 (2025)
3. Fernandez, A., Herrera, F., Cordon, O., del Jesus, M.J., Marcelloni, F.: Evolutionary fuzzy systems for explainable artificial intelligence: Why, when, what for, and where to? IEEE Comput. Intell. Mag. 14(1), 69–81 (2019)
4. Márquez, F.A., Peregrín, A., Herrera, F.: Cooperative evolutionary learning of fuzzy rules and parametric aggregation connectors for Mamdani linguistic fuzzy systems. IEEE Trans. Fuzzy Syst. 15(6), 1162–1178 (2007)
5. Martín-Moreno, J., Márquez, F.A., Roldán, A.M., Peregrín, A.: Online on-device adaptation of linguistic fuzzy models for TinyML systems. AI 6(12), 325 (2025)
6. Angelov, P.P., Filev, D.P.: An approach to online identification of Takagi-Sugeno fuzzy models. IEEE Trans. Syst., Man, Cybern. B, Cybern. 34(1), 484–498 (2004)
7. Angelov, P., Buswell, R.: Evolving rule-based models: a tool for intelligent adaption. In: Proc. IFSA World Congr. 20th NAFIPS Int. Conf., vol. 3, pp. 1062–1067 (2001)
8. Angelov, P.: Fuzzily connected multimodel systems evolving autonomously from data streams. IEEE Trans. Syst., Man, Cybern. B, Cybern. 41(4), 898–910 (2011)
9. Ge, D., Zeng, X.J.: Learning evolving T–S fuzzy systems with both local and global accuracy – a local online optimization approach. Applied Soft Computing 68, 795–810 (2018).
10. Angelov, P.P., Gu, X., Principe, J.C.: Autonomous learning multimodel systems from data streams. IEEE Trans. Fuzzy Syst. 26(4), 2213–2224 (2018)
11. Ge, D., Zeng, X.J.: A self-evolving fuzzy system which learns dynamic threshold parameter by itself. IEEE Trans. Fuzzy Syst. 27(8), 1625–1637 (2019)
12. Ge, D., Zeng, X.J.: Learning data streams online - an evolving fuzzy system approach with self-learning/adaptive thresholds. Inf. Sci. 507, 172–184 (2020)
13. Kasabov, N.: Evolving fuzzy neural networks for supervised/unsupervised online knowledge-based learning. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 31(6), 902–918 (2001).
14. Kasabov, N.K., Song, Q.: DENFIS: dynamic evolving neural-fuzzy inference system and its application for time-series prediction. IEEE Trans. Fuzzy Syst. 10(2), 144–154 (2002)
15. Pratama, M., Anavatti, S.G., Angelov, P.P., Lughofer, E.: PANFIS: A novel incremental learning machine. IEEE Transactions on Neural Networks and Learning Systems 25(1), 55–68 (2013).
16. Pratama, M., Anavatti, S.G., Lughofer, E.: GENEFIS: Toward an effective localist network. IEEE Transactions on Fuzzy Systems 22(3), 547–562 (2013).