

## A Novel Fuzzy Learning Model based on Forgetting Factor

Sajad Haghzad Klidbary<sup>1</sup>,

1- Faculty of Electrical and Computer Engineering, University of Zanjan, Zanjan, Iran.

<sup>1</sup>s.haghzad@znu.ac.ir

Corresponding author's address: Sajad Haghzad Klidbary, Faculty of Electrical and Computer Engineering, University of Zanjan, Zanjan, Iran.

**Abstract-** Biological observations, indicate that amnesia is an integral part of the human learning system. Thus, amnesia in learning algorithms is not necessarily destructive and can be constructive. In implementation, due to space constraints and the number of neurons, a limited number of training patterns can be taught to the network. Consequently, to be able to obtain long-term learning capability, it must possess a kind of forgetting mechanism to make space for new learning patterns. Thus, a type of forgetting mechanism similar to the function of the human brain is necessitated. The need for a forgetting mechanism is more acute in online training. Amnesia is modeled as the loss of information from memory. In this paper, the ALM, which is one of the most widely used methods, is employed. The selected algorithm models the system based on the distribution of ink drops based on training data. In this method, in all the implementations, the amplitude of the ink drops on the IDS planes remains unchanged, and no amnesia occurs, which is contrary to biological observations. In this work, the forgetting mechanism is added to the presented algorithm, and simulations in the modeling process are investigated.

**Keywords-** Active Learning Method (ALM), Ink Drop Spread (IDS) Operator, Fuzzy Inference System, Artificial Neural Network, Forgetting Factor, Modelling.

## ارائه یک مدل یادگیری فازی جدید مبتنی بر معیار فراموشی

سجاد حق زاد کلیدبری<sup>۱</sup>

۱- دانشکده مهندسی برق و کامپیوتر، دانشگاه زنجان، زنجان، ایران.

<sup>1</sup> s.haghzaad@znu.ac.ir

\* نشانی نویسنده مسئول: سجاد حق زاد کلیدبری - ایران - زنجان - کیلومتر ۶ جاده تبریز - دانشگاه زنجان - دانشکده مهندسی.

چکیده- مشاهدات بیولوژیکی بیان می‌دارد که فراموشی، جزء جداناپذیر از سیستم یادگیری انسان است. بنابراین فراموشی در الگوریتم‌های یادگیری لزوماً مخرب نبوده و می‌تواند سازنده نیز باشد. در پیاده‌سازی‌ها، به دلیل محدودیت فضا و تعداد نورون‌های شبکه، تعداد محدودی الگوی آموزشی قابل آموزش بوده و الگوهای بعدی با این الگوها تداخل مخرب پیدا خواهند کرد؛ در نتیجه، الگوریتم‌ها، برای یادگیری درازمدت، باید نوعی مکانیزم فراموشی داشته باشند تا فضای یادگیری (ذخیره سازی) برای الگوهای آموزشی جدید ایجاد گردد. بنابراین، برای موفقیت در حوزه یادگیری ماشین، نیازمند نوعی مکانیزم فراموشی مشابه عملکرد مغز انسان هستیم. فراموشی به صورت از دست رفتن اطلاعات از حافظه‌ها مدل می‌شود و لزوم وجود این مکانیزم، در آموزش آنلاین محسوس‌تر است چراکه شبکه باید دائماً وزن‌های خود را بروز کند. در این مقاله از روش یادگیری فعال که یکی از روش‌های پرکاربرد می‌باشد، بهره گرفته شده است. این روش بر مبنای پخش قطرات جوهر به‌ازای داده‌های آموزشی به مدل‌سازی سیستم می‌پردازد. در این روش، دامنه قطرات جوهر بر روی صفحات بدون تغییر مانده و هیچ‌گونه فراموشی صورت نمی‌پذیرد که مغایر با مشاهدات بیولوژیکی است. در این مقاله مکانیزم فراموشی به این الگوریتم اضافه شده و شبیه‌سازی‌ها نشان از افزایش قدرت محاسباتی مدل پیشنهادی در برخورد با مجموعه داده‌های متفاوت دارد.

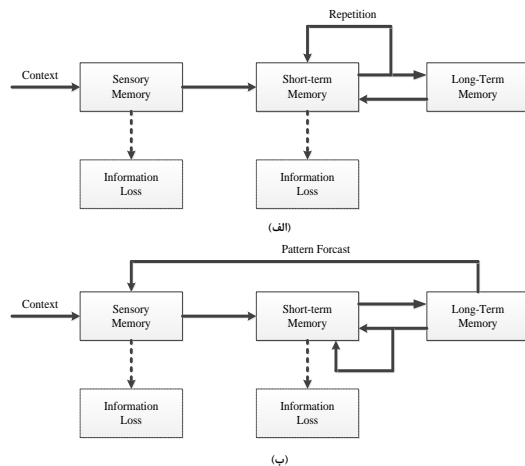
واژه‌های کلیدی: روش یادگیری فعال، اپراتور پخش قطره جوهر، سیستم استنتاج فازی، شبکه‌های عصبی بیولوژیکی، معیار فراموشی، مدل‌سازی.

### ۱- مقدمه

نورون برحسب میزان و زمان‌بندی فعالیت‌های آن دو نورون نسبت به یکدیگر تعیین می‌گردد و تحت نام قانون هب نام برده می‌شود. زمانی که پتانسیل غشاً به مقدار ترشولد رسید، یک سیگنال در غشاً ایجاد می‌گردد که با حرکت این سیگنال و رسیدن آن به انتهای آکسون، آزاد شدن یون‌ها در فاصله سیناپسی را در پی دارد. بنابراین ضریب سیناپسی را می‌توان قدرت باز یا بسته شدن درچه‌ها برای ورود یا خروج یون‌ها دانست. طبق قانون هب، زمانی در یک فاصله سیناپسی هدایت بهتر می‌شود، یعنی اینکه دائماً سیگنال از آن‌ها عبور می‌کند. بنابراین طبق این قانون، مکان‌هایی که سیگنال بیشتر منتقل می‌شود، ضریب سیناپسی قوی‌تر است، چون یون‌های بیشتری در آن قسمت حضور دارند [۴، ۵]. بنابراین

این امروزه حوزه‌های علمی مختلفی نظیر شبکه‌های عصبی مصنوعی و منطق فازی را می‌توان یافت که به بررسی، بازسازی و ارائه الگوریتم نحوه پردازش اطلاعات مغز انسان می‌پردازد [۱، ۲]. یکی از چالش‌های پیش‌روی الگوریتم‌های یادگیری، مکانیزم فراموشی است. همان‌طور که می‌دانیم، فراموشی جز جدانشدنی از سیستم یادگیری انسان می‌باشد. هب روانشناس کانادایی بیان نمود که اگر فعال شدن نورون A سبب فعال شدن نورون B شود با توجه به علی بودن سیستم ضریب سیناپسی بین نورون‌ها باید تقویت گردد [۳]. ضمناً نشان داد که شدت اتصال سیناپسی بین دو

فراموشی را انجام دهد تا فضا برای الگوهای آموزشی جدید ایجاد شود. بدیهی است که پارامترهای غیرضروری باید برای فراموشی در اولویت باشند. این لزوم وجود مکانیزم فراموشی، در آموزش آنلاین محسوس تر است چراکه شبکه باید دائماً وزنهای خود را بروز کند. همان طور که در مدل های ارائه شده، قابل مشاهده است، فراموشی به صورت از دست رفتن (نشت کردن) اطلاعات از حافظه ها مدل شده است.



شکل ۱ (الف) مدل پردازش سه گانه در ذهن انسان (ب) مدل حافظه بر اساس تابع پیش بینی برای پردازش در ذهن انسان [۶].

الهام از طبیعت و به ویژه توجه به ساختار داخلی مغز همواره مورد توجه محققان بوده است. سیستم های الهام گرفته شده از بیولوژیک، سیستم هایی هستند که از رفتارهای مشاهده شده در طبیعت الهام گرفته و تلاش در بهره گیری از آنها در انجام محاسبات را دارند. بنابراین می توان بیان داشت که فراموشی، جز جداناپذیر از سیستم های یادگیری موجود است. حافظه دار بودن و در کنار آن داشتن مکانیزم فراموشی در یک سیستم یادگیری در کاربردهای که با پدیده های متغیر با زمان روبرو هستیم از اهمیت زیادی برخوردار است. به عنوان مثال در یک کاربرد کنترلی که سیستم با زمان در حال تغییر است، دنبال کردن تغییرات سیستم از اهمیتی زیادی برخوردار است.

در این مقاله به عنوان پایه مدل پیشنهادی، از الگوریتم ALM بهره گرفته شده و با افزودن معیار فراموشی به صفحات IDS سعی در افزودن مکانیزم فراموشی به الگوریتم ALM داریم. در این راستا، همانند فضای یونی مغز، که تغییرات چگالی یونها اطراف نورون هدف، سبب تاثیرگذاری روی نورون های اطراف شده و روی ضرایب سیناپسی تاثیرگذار است [۴]. در مدل پیشنهادی، پخش قطره جوهر راحت تر، معادل ضرایب سیناپسی بزرگ تر و در نتیجه آتش شدن تعداد نورون های همسایه بیشتر است. با فرض معادل گرفتن، هر یک از پیکسل ها صفحات IDS در روش یادگیری فعال

در محیط الکتروشیمیایی مغز انسان، اگر یون های بعدی با فاصله زمانی مناسبی به محیط سوما نرسند، نورون آتش نخواهد کرد (بنابراین طبق نظریه هب به نوعی فراموشی رخ داده است).

طبق مشاهدات بیولوژیکی، برای مدل سازی عملکرد مغز انسان، از مدل سازی جبری "حافظه" و "یادگیری" استفاده شده و یک چارچوب کلی برای عملکرد مجموعه ذهن انسان معرفی شده است [۶]. طبق تعاریف، ذهن انسان دارای سه قسمت حافظه، به نام های حافظه حسی، حافظه کوتاه مدت و حافظه بلندمدت است. بلوک دیاگرام مدل سه گانه حافظه در شکل ۱-الف آورده شده است. حافظه حسی<sup>۱</sup>، اولین حافظه ای است که اطلاعات به آن وارد می شود و البته پردازشی روی داده ها انجام نمی دهد (در مرحله شناسایی محرک). پس از عبور از حافظه حسی، اطلاعات وارد حافظه کوتاه مدت<sup>۲</sup> می شود (گزینش این اطلاعات ارسالی، براساس مرتبط بودن آنها به تکالیفی است که باید انجام گیرد و بنابراین توجه خاص، اطلاعات را به حافظه کوتاه مدت هدایت می کند و یا به بیانی دیگر از نظر مدل سازی، این اطلاعات وزن دهی می شوند). اطلاعات می توانند تا مدت حدود یک دقیقه در حافظه کوتاه مدت باقی بمانند و سپس وارد حافظه بلند مدت<sup>۳</sup> می شوند که در آن بخش قابلیت باقی ماندن به مدت نامحدود را دارند. اطلاعات به وسیله پردازش کنترل شده ای که در حافظه کوتاه مدت صورت می گیرد، در حافظه دراز مدت نگهداری می شود (پردازش کنترل شده، مانند تجربه، تمرین و متصل کردن اطلاعات به اطلاعات دیگر است). بنابراین نگهداری در حافظه بلندمدت مستلزم تلاش است و این اطلاعات مانند اطلاعات مربوط به آموزش دوچرخه سواری یا پرتاب توپ است که در حافظه ثبت می شوند و هرگز فراموش نمی شوند. زمانی که اطلاعات به طریقی از حافظه کوتاه مدت به حافظه بلندمدت منتقل شوند، به اصطلاح، آموختن صورت گرفته است. طبق مدل بیان شده برای ذهن، اطلاعات ذخیره شده در حافظه بلندمدت، برای پردازش شدن باید دوباره وارد حافظه کوتاه مدت شوند. طبق تحقیقات صورت گرفته از سیگنال های عصبی، مدل کامل تری از مدل قبلی ارائه شده است که بلوک دیاگرام آن در شکل ۱-ب دیده می شود.

همان طور که در مدل های فوق قابل مشاهده است، فراموشی جزئی جدانشدنی از یادگیری محسوب می شود. نکته ای دیگر که باید به آن توجه نمود این است که، در پیاده سازی عملی، به دلیل محدودیت فضا و تعداد نورون های شبکه، تعداد محدودی الگوی آموزشی قابل آموزش به شبکه است و الگوهای بعدی با این الگوها تداخل مخرب پیدا خواهند کرد؛ در نتیجه، برای این که مغز مصنوعی قادر به یادگیری درازمدت باشد بایستی بتواند نوعی عمل

که وزن‌های شبکه، از حالت بهینه برای پاسخ به دیتاست اول خارج شوند؛ یعنی پاسخ شبکه به دیتاست‌های آموزشی قبلی بدتر می‌شود[۸]. پدیده فراموشی مخرب، یک پدیده ذاتی در آموزش شبکه‌های عصبی مصنوعی است.

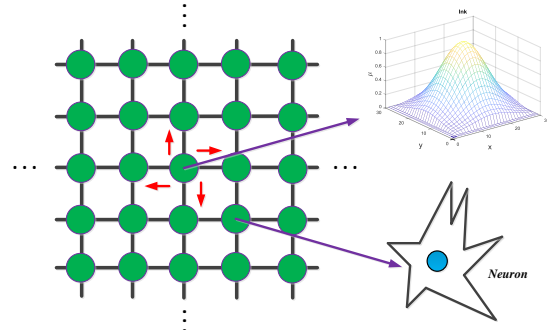
## ۲-۲- فراموشی هدفمند

این فراموشی در الگوریتم‌های یادگیری ماشین معمولاً برای ساده کردن قواعد تصمیم‌گیری و نیز حذف داده‌های غیرضروری یا مضر استفاده می‌شود. در مسائل طبقه‌بندی، اگر قواعد تصمیم‌گیری به صورت مداوم آموزش داده شود، ممکن است خیلی پیچیده شود؛ بنابراین نوعی مکانیزم برای حذف اثر آموزش‌های کم‌اهمیت قبلی لازم است. مکانیزم حذف یا "فراموشی" می‌تواند تنها تابع زمان باشد، به طوری که مقدار یک داده به صورت  $c = ke^{-at}$  کاهش یابد، مگر این‌که این داده در صورت ضرورت، توسط الگوریتم تقویت گردد. به این نوع فراموشی که تنها تابع زمان است، فراموشی غیرفعال می‌گویند. با این حال، برای مؤثرتر شدن اثر فراموشی، می‌توان داده‌هایی را که برای تصمیم‌گیری صحیح مضر هستند پیدا کرد و آن‌ها را سریع‌تر حذف (فراموش) کرد، که به این عمل فراموشی فعال گفته می‌شود. فرمول کلی فرآیند فراموشی وزن‌ها در آموزش وزن‌های شبکه عصبی با استفاده از ضریب فراموشی صورت می‌گیرد که این ضریب سرعت فراموشی وزن‌ها را تعیین می‌نماید[۹]. این ضریب فراموشی می‌تواند متغیر باشد و در طول فرآیند آموزش به صورت خود تطبیقی تغییر نماید. بدیهی است که رابطه فراموشی فوق، مستقل از طرز محاسبه  $\Delta W$  بوده و لذا به تمام الگوریتم‌های آموزشی (گرادیان کاهشی، نیوتن و موارد مشابه)، اعم از آنلین یا آفلاین، قابل اعمال است. به عنوان مثال می‌توان در الگوریتم‌هایی نظیر MLP بیان داشت که نرخ یادگیری عمل فراموشی را انجام می‌دهد.

## ۲-۳- فراموشی در الگوریتم‌های حافظه محور

در روش‌های حافظه‌محور، خروجی شبکه بر اساس شباهت ورودی با ورودی‌های پیشین که در حافظه ذخیره شده‌اند، صورت می‌پذیرد. مشابه آنچه در شبکه‌های عصبی RBF و الگوریتم KNN انجام می‌گیرد و خروجی این الگوریتم‌ها به صورت تابعی از فاصله ورودی از مقادیر ذخیره شده محاسبه می‌شود. الگوریتم‌های MBL مستلزم ذخیره‌سازی و در دسترس بودن تمامی جفت‌های ورودی/خروجی قبلی در فرآیند یادگیری هستند و همین امر سبب افزایش پیچیدگی محاسباتی با افزایش تعداد داده‌های آموزشی خواهد بود. در این روش‌ها امکان تطبیق شبکه با داده‌های آموزشی جدید و دور ریختن (پاک کردن) آموزش‌های قدیمی از

معادل با یک نورون، پخش قطره جوهر در صفحات IDS معادل با تغییرات چگالی یونی در اطراف نورون هدف بوده که هر چقدر ضرایب سیناپسی اطراف نورون هدف بزرگتر باشد، تغییرات یونی در آن مسیرها بیشتر خواهد بود. در شکل ۲) نمایشی از عملکرد بیان شده آورده شده است.



شکل ۲) نمایش یک شبکه نورونی، با مجموعه‌ای از نورون‌ها، همچنین نمایشی از پخش قطره جوهر گاوسی که همانند پخش یونی در نورون هدف صورت می‌پذیرد.

ساختار مقاله در ادامه به شرح زیر است: در بخش دوم انواع مکانیزم‌های فراموشی مورد بررسی قرار می‌گیرد. در بخش سوم مفاهیم روش یادگیری فعال و اپراتور پخش قطره جوهر ارائه شده است. روش ارائه شده در بخش چهارم بررسی می‌شود. در بخش پنجم شبیه‌سازی‌ها و ارزیابی الگوریتم ارائه شده آورده شده است. در نهایت بخش ششم شامل جمع‌بندی، نتیجه‌گیری و پیشنهادات برای کارهای آینده است.

## ۲- انواع مکانیزم‌های فراموشی

با توجه مشاهدات بیولوژیکی و اهمیت موضوع فراموشی، الگوریتم‌های مختلفی با بهره‌گیری از مکانیزم‌های متفاوت فراموشی ارائه شده است. دسته‌بندی فراموشی در روش‌های آموزشی موجود در ادامه آورده شده و مورد بررسی قرار گرفته شده است.

### ۲-۱- فراموشی مخرب

این فراموشی، عبارت است از بین رفتن اثر آموزش‌های قبلی روی وزن‌های شبکه عصبی در اثر آموزش‌های جدید[۷]. در واقع، در اثر آموزش شبکه با یک مجموعه داده آموزشی، وزن‌های شبکه طوری تنظیم می‌شوند که پاسخ مناسبی برای آن دیتاست تولید شود. وقتی دیتاست بعدی به شبکه آموزش داده شود، وزن‌ها دوباره تغییر می‌کنند تا شبکه بتواند برای دیتاست دوم نیز پاسخی تا حد امکان صحیح تولید کند. این تغییر وزن‌ها موجب می‌شود

دریافتی، مجموعه تصمیمات صحیح را پیدا می‌کند. در آموزش شبکه‌های عصبی با روش‌های پاداش‌محور نیز تغییرات وزن‌ها می‌تواند به فراموشی مخرب بینجامد. فراموشی در ذات الگوریتم‌های پاداش‌محور نیز وجود دارد و ماشین با دریافت پاداش بیشتر به-ازای یک تصمیم خاص در یک حالت معین، آن تصمیم را در آن حالت در اولویت قرار می‌دهد و تصمیمات دیگر را فراموش می‌کند [۱۳]. همچنین فراموشی در چنین الگوریتم‌هایی می‌تواند به‌سادگی به‌صورت از بین رفتن ارزش‌های آموخته‌شده و پاداش‌های قبلی و استفاده نشدن آن‌ها در اتخاذ تصمیمات جدید تعریف شود. این روند فراموشی می‌تواند باعث شود که الگوریتم همگرا نشود. از طرفی، اگر شرایط محیط متغیر باشد، وجود این فراموشی برای تطبیق‌پذیری ماشین با محیط لازم است.

بنابراین با معرفی روش‌های مرتبط با فراموشی می‌توان بیان داشت که فراموشی، جزء جداناپذیر از سیستم‌های یادگیری موجود است. حافظه‌دار بودن و داشتن مکانیزم فراموشی در یک سیستم یادگیری در کاربردهای که با پدیده‌های متغیر با زمان روبرو هستیم از اهمیت زیادی برخوردار است. مثلاً در یک کاربرد کنترلی متغیر با زمان، دنبال کردن تغییرات سیستم از اهمیتی زیادی برخوردار است. در ادامه به بررسی روش ALM و افزودن مکانیزم فراموشی به آن آورده است.

### ۳- روش یادگیری فعال (ALM)

این روش در سال ۱۹۹۷ ارائه شده است و تاکنون توسعه یافته-های بسیاری از آن ارائه شده است [۱۴]. این الگوریتم و توسعه یافته‌های آن توانسته نتایج خوبی را در حوزه‌های مختلفی کنترل، مدل‌سازی، کلاس‌بندی، خوشه‌بندی و پردازش تصویر از خود نشان دهد [۱۵-۱۹]، [۲۰-۲۳]، [۲۴، ۲۵]. عملکرد این روش به این نحو است که در برخورد با مسائل پیچیده، ابتدا مسئله را به مسائل کوچکتر می‌شکند و روی نتایج حاصل از این زیرسیستم‌ها، برهم-نهی انجام داده و خروجی نهایی بدست آورده می‌شود. این الگوریتم در اولین گام یک سیستم چند-ورودی تک-خروجی را به چند زیرسیستم تک-ورودی تک-خروجی می‌شکند. مکانیزم عملکرد این الگوریتم در شکل ۳-الف) نمایش داده شده است.

### ۳-۱- اپراتور IDS

روش یادگیری فعال برخلاف بسیاری از الگوریتم‌های مرسوم بجای بهره‌گیری از روابط ریاضی پیچیده و بیان مسائل بصورت عددی در دنیای کلاسیک، با بیان کیفی، سعی در مدل‌سازی عملکردی ذهن انسان دارد.

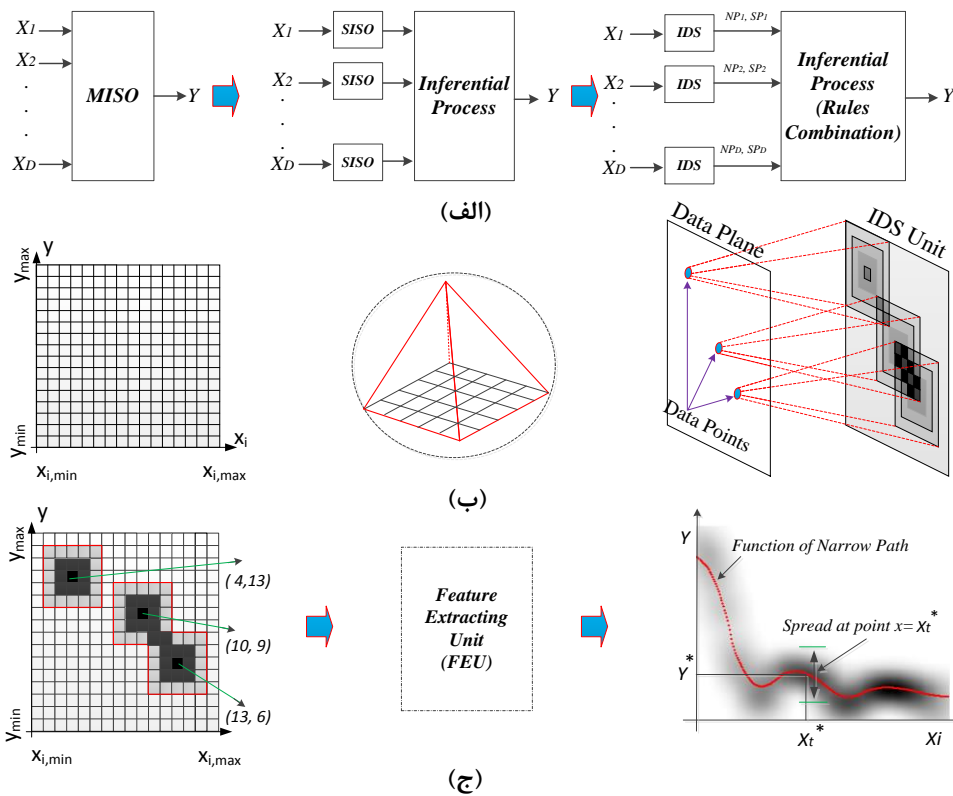
حافظه وجود دارد. در نتیجه برای ایجاد مطابقت با داده ورودی، در روش‌های MBL، باید الگوریتمی برای حذف نمونه‌های غیرکارآمد که دیگر نشان دهنده نگاشت کنونی شبکه نیستند، در نظر گرفته شود. به چنین الگوریتم‌هایی، "الگوریتم‌های کاهش نمونه" گفته می‌شود. این الگوریتم‌ها، داده‌های آموزشی قدیمی که دیگر برای تصمیم‌گیری به کار نمی‌آیند، یافته و از حافظه حذف می‌کند. این الگوریتم‌های فراموشی، باعث انعطاف و تطبیق‌پذیری بیشتر الگوریتم حافظه‌محور می‌شوند. الگوریتم‌های فراموشی مبتنی بر زمان و فراموشی مبتنی بر مکان از جمله این الگوریتم‌های فراموشی هستند [۱۰]. در الگوریتم‌های نوع اول، به نمونه‌های داخل حافظه یک وزن تخصیص داده می‌شود و در زمان ورود داده این وزن مقدار یک دارد و با گذشت گام‌های زمانی، این وزن‌ها برورزسانی می‌شوند. در الگوریتم‌های نوع دوم، آموخته‌های قبلی بر اساس نزدیکی به مشاهدات جدید فراموش می‌شوند و وزن‌های تخصیص داده شده به داده‌ها با توجه به مشاهدات بعدی و میزان شباهت با آن‌ها، کاهش می‌یابند.

### ۲-۴- فراموشی در شبکه‌های اسپایکی

فراموشی در شبکه‌های اسپایکی نیز مشابه فراموشی در شبکه‌های عصبی عادی است و می‌تواند به همان فرم کاهش وزن‌ها بمرور زمان باشد. وزن‌ها با توجه به اسپایک‌های وارد شده به آن‌ها، تا مدت‌زمان معینی افزایش و سپس در مدت‌زمان معینی کاهش می‌یابند. هر چه یک وزن خاص در شبکه پرکاربردتر باشد، اسپایک‌های بیشتری وارد آن می‌شوند. در نتیجه، برای این که وزن‌های با اهمیت‌تر مقدار بیشتری پیدا کنند، باید نرخ افزایش وزن در مدت‌زمان افزایش، با تعداد اسپایک‌ها نسبت مستقیم و در مدت زمان کاهش، با تعداد اسپایک‌ها نسبت عکس داشته باشد. نوعی دیگری از فراموشی نیز به‌طور طبیعی در شبکه‌های عصبی اسپایکی وجود دارد که مربوط به کاهش تدریجی پتانسیل‌های غشایی در طول زمان است. با ورود هر اسپایک به یک نورون، پتانسیل غشایی آن ابتدا برای مدت معینی افزایش می‌یابد اما پس از آن شروع به کم شدن می‌کند و اگر اسپایک دیگری وارد نورون نشود، پتانسیل غشایی آن در نهایت به صفر خواهد رسید [۳، ۱۱، ۱۲].

### ۲-۵- فراموشی در روش‌های آموزشی تقویتی (پاداش محور)

در الگوریتم‌های پاداش محور، ماشین یا شبکه تحت آموزش در هر حالت خود، با مجموعه‌ای از تصمیمات ممکن روبرو می‌شود و با انتخاب هر تصمیم، پاداش متناظر با آن را دریافت می‌کند. با تکرار الگوریتم، شبکه در مسیر تلاش برای بیشینه کردن پاداش



شکل ۳ (الف) عملکرد روش یادگیری فعال در برخورد با سیستم‌های چند- ورودی تک- خروجی، (ب) صفحه مشبک خالی از قطره جوهر و قطره جوهر هرمی، (ج) اعمال داده‌های آموزشی و واحد استخراج کننده ویژگی‌ها.

موتور اصلی روش یادگیری فعال اپراتور پخش قطره جوهر است که عدم قطعیت را مدل‌سازی می‌کند. هر کدام از زیر سیستم‌های تک- ورودی تک- خروجی توسط صفحات IDS مدل می‌شود. در صورت کوانتیزه کردن فضای ورودی و خروجی صفحات مشبکی از نقاط داده تصویر شده می‌باشند. اثر IDS روی مجموعه نقاط تجربه همانند چکاندن یک قطره جوهر روی هر نقطه در صفحه است. این فرآیند در شکل ۳-ب آورده شده است. بعد از اعمال تمامی داده‌های آموزشی، همانند شکل ۳-ج، شکل قطرات جوهر با توابع عضویت هرمی در برخی نقاط با هم، هم‌پوشانی داشته و نواحی پیوسته‌ای با درجه تعلق‌های مختلفی را ایجاد نموده‌اند. بعد از برهم نهی تمامی داده‌های آموزشی، الگویی ایجاد شده و دو ویژگی مسیر باریک<sup>۵</sup> و پراکندگی<sup>۶</sup> استخراج می‌شود. مسیر باریک رابطه بین ورودی و خروجی سیستم را بیان می‌دارد و شعاع پراکندگی بیانگر اهمیت متغیر ورودی  $x_i$  در مقایسه با سایر متغیرهای ورودی در تعیین مقدار خروجی سیستم است.

و با فرض بعد  $D$ ، برای هر داده آموزشی به صورت زیر داریم:

$$x_i = (x_{i1}, x_{i2}, \dots, x_{iD})^T, x_i \subseteq R^D \quad (2)$$

اولین گام در روش ALM، کوانتیزه نمودن متغیرهای ورودی و خروجی است. با فرض سطوح کوانتیزاسیون به ترتیب برابر با  $R_{sn_x}$  و  $R_{sn_y}$ ، که معمولاً در شبیه‌سازی‌ها این مقادیر برابر در نظر گرفته می‌شود، داریم:

$$(x_{i1}, x_{i2}, \dots, x_{iD})^T \xrightarrow{\text{Quantization}} X_{iq} \in \{1, 2, \dots, R_{sn_x}\}, \quad (3)$$

$$(y_1, y_2, \dots, y_N)^T \xrightarrow{\text{Quantization}} Y_q \in \{1, 2, \dots, R_{sn_y}\}. \quad (4)$$

اگر فرض کنیم  $p(x, y)$  یک نقطه در صفحه  $y - x_i$  باشد و تیرگی در این نقطه از صفحه با  $d(x, y)$  بیان گردد و قطره جوهر همانند یک تابع گاوسی با واریانس مناسب باشد که در مرکز این تابع گاوسی مقدار برابر با یک است. در این صورت برورسانی صفحه IDS پس از اعمال داده آموزشی جدید در نقطه  $(x_s, y_s)$  به صورت زیر قابل بیان خواهد بود.

$$P_{x_i y} = \{p(x, y) | x \in X_i, y \in Y\}, \quad (5)$$

$$d(x_s + u, y_s + v) = d(x_s, y_s) + h(u, v), \quad (6)$$

که در رابطه فوق  $R$  بیانگر شعاع قطره جوهر و  $h$  شکل تابع قطره جوهر است (در اکثر شبیه‌سازی‌ها، قطره جوهر یک تابع گاوسی انتخاب می‌گردد). با استفاده از تعاریف بیان شده می‌توان توابع  $\psi$

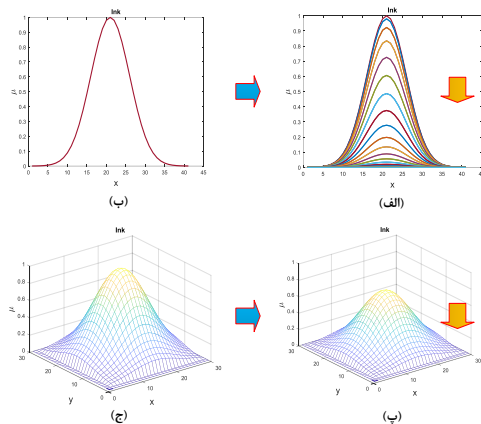
روش‌های مختلفی برای محاسبه مسیر باریک و پراکندگی وجود دارد که در ادامه به بررسی روش بیان شده در مقاله [۲۶] پرداخته شده است. با فرض مجموعه  $S$  که شامل  $N$  داده آموزشی است:

$$S = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}, \quad (1)$$

شدن می‌کند و اگر اسپایک دیگری وارد نورون نشود، پتانسیل غشایی آن در نهایت به صفر خواهد رسید.

این بیان را می‌توان معادل پخش قطرات جوهر دانست و مابه-ازای تکرار قطرات جوهر در فضای یونی مغز را می‌توان، تجمیع یون‌ها در بخش سوما پنداشت که اگر یون‌های بعدی به‌موقع نرسد تأثیر دیرکرد در آتش‌شدن نورون را به همراه دارد. بنابراین معادل این بیان را در روش یادگیری فعال، می‌توان با همان تکرار تجربه و تأیید قطرات و پخش مجدد قطرات جوهر در یک نقطه دانست.

پارامتر فراموشی باید در سیستم‌های محاسباتی مورد توجه قرار گیرد و با اعمال برچسب زمانی (فاکتور زمان) روی داده‌های ورودی آموزشی، داده‌هایی که دیرتر آموزش دیده‌اند نقش بیشتری در فرآیند تصمیم‌گیری نهایی داشته باشند. در راستای دستیابی به این هدف با اضافه نمودن یک ضریب فراموشی به روش یادگیری فعال تغییرات زمانی اعمال شده است. بنابراین مکانیزم فراموشی را می‌توان در صفحات IDS در نظر گرفت و بسته به مقدار پارامتر، تأثیرگذاری آن‌ها قابل‌رؤیت خواهد بود. مکانیزم فراموشی ناشی از فرآیند دیفیوژن و وجود دینامیک در فرآیند دیفیوژن (کاهش ارتفاع قطره جوهر با گذشت زمان و پارامترهایی مثل سرعت پخش جوهر و تکرار تجربه) رخ می‌دهد. تاکنون در تمام پیاده‌سازی‌های موجود، قطرات جوهر بر روی صفحات پخش قطرات ثابت مانده و هیچ‌گونه فراموشی رخ نمی‌داد، که این امر مغایر با بیولوژیک است. مکانیزم فراموشی به صفحات IDS افزوده شده است تا الگوریتم مشابهت بیشتری با بیولوژیک پیدا کند. بنابراین قطرات جوهر از یک ارتفاعی مشخص شروع به پخش می‌کنند و آرام‌آرام ارتفاع آن‌ها کاهش می‌یابد. کاهش ارتفاع قطره جوهر بیانگر این است که درجه باور نسبت به نقاط مشاهده باگذشت زمان و عدم تکرار مشاهده کاهش پیدا می‌کند و این نمودی از فراموشی است. برای درک بهتر، مثالی از این روند در شکل (۴) آورده شده است.



شکل (۴) تغییرات قطره جوهر با گذشت زمان در حالت دو بعدی و سه بعدی، کاهش ارتفاع قطره جوهر باگذشت زمان (کاهش درجه تعلق).

و S را به ترتیب برای یافتن مسیر باریک و پراکندگی در صفحه IDS تعریف نمود [۲۶]:

$$\psi_{x_i}(x) = \left\{ b \mid \sum_{y=y_{min}}^b d(x,y) \approx \sum_{y=b}^{y_{max}} d(x,y), b \in Y \right\}, \quad (7)$$

$$S_{x_i}(x) = \max_{y \in Y} \{y \mid d(x,y) > T\} - \min_{y \in Y} \{y \mid d(x,y) > T\}, \quad (8)$$

رابطه اول بیان می‌دارد که مقدار تابع مسیر باریک به‌ازای هر ورودی تست x برابر با b است به شرطی که مجموع جوهرهای ذخیره شده در سلول‌های بالای مختصات (x, b) در صفحه دو بعدی با یک تلورانس (تقریبی) برابر با مجموع مقادیر جوهرهای ذخیره‌شده در سلول‌های پایین مختصات (x, b) باشد. رابطه دوم بیان می‌دارد که مقدار پراکندگی (شعاع پخش) در نقطه‌ای مانند x متناسب با عرض موثر الگوی قرار گرفته بر روی صفحه IDS در ستون متناظر با x است. در این رابطه پارامتر T بیانگر مقدار آستانه برای یافتن و تأثیرگذاری مقدار تیرگی ذخیره شده در سلول برای مقدار پراکندگی است.

### ۳-۲- استنتاج در روش یادگیری فعال

با آمدن داده جدید  $x = (x_1, x_2, \dots, x_D)$  که D بیانگر بعد داده ورودی است، مقادیر مسیر باریک و شعاع پراکندگی در نقطه x توسط توابع فوق (روابط ۷ و ۸) محاسبه شده و در نهایت روش ALM بر مبنای مقادیر فوق برای یافتن مقدار خروجی طبق روابط ۹ و ۱۰ برای بخش استنتاج دارد:

$$\beta_i \propto \frac{1}{S_{x_i}(x_i)}, \quad (9)$$

$$y(x) = \sum_{i=1}^c \beta_i \times \psi_{x_i}(x_i), \quad (10)$$

بنابراین خروجی روش یادگیری فعال جمع وزن‌داری از خروجی تک تک زیرسیستم‌ها است.

### ۴- الگوریتم پیشنهادی

در سیستم یادگیری انسان همیشه فراموشی وجود داشته و اطلاعات با گذشت زمان فراموش شده و همچنین برخی از اطلاعات ممکن است که به‌طور کامل از ذهن انسان فراموش شوند. طبق مشاهدات بیولوژیکی، فراموشی در شبکه‌های عصبی وجود دارد که مربوط به کاهش تدریجی پتانسیل‌های غشایی در طول زمان است. با ورود هر اسپایک به یک نورون، پتانسیل غشایی آن ابتدا برای مدت معینی افزایش می‌یابد اما پس از آن شروع به کم

محیط فیدبک گرفته شده و این فیدبک رو در بخش استنتاج سیستم تاثیرگذار است. روش پیشنهادی برخلاف روش هب نیازمند روش‌های پیچیده محاسباتی نیست و در جاهایی که قطره بیشتر چکانده شده، فراموشی کمتر و ضریب سیناپسی در پیاده-سازی اسپایکی قوی‌تر خواهد بود. ضمن اینکه، در بسیاری از الگوریتم‌های موجود، با ورود یک داده جدید باید یادگیری روی کل وزن‌های شبکه اجراء شده و بروزسانی صورت بگیرد که علاوه بر پیچیدگی زمانی، نیازمند زمان محاسباتی زیادی نیز است.

## ۵- شبیه سازی

در این بخش جهت ارزیابی الگوریتم ارائه شده شبیه‌سازی‌هایی به‌ازای ورودی‌های مختلف صورت پذیرفته است. تمامی شبیه-سازی‌ها در برنامه متلب 2017a و با یک کامپیوتر شخصی با پردازنده Core i5، 2.4 GHz و 4 GB رم صورت گرفته شده است. در شبیه‌سازی‌ها تعداد سطوح کوانتیزاسیون محورها برابر با ۲۵۶ در نظر گرفته شده است. برای مقایسه مدل پیشنهادی به مقایسه آن با الگوریتم‌های MLP و ANFIS پرداخته شده است.

## ۵-۱- طبقه بندی

در ابتدا به بررسی در فضای طبقه‌بندی پرداخته شده است. در جدول ۱ مشخصات و ویژگی‌های انواع ست‌داده‌های بهره گرفته شده برای شبیه‌سازی آورده شده است. روش ارزیابی 10-Fold در تمامی شبیه‌سازی‌ها استفاده شده است.

جدول ۱: مشخصات مجموعه داده‌های استفاده شده برای ارزیابی.

منبع	کلاس	ویژگی	نمونه	ست داده
<a href="http://sci2s.ugr.es/keel/category.php?cat=clas#sub2">http://sci2s.ugr.es/keel/category.php?cat=clas#sub2</a>	2	2	5300	Banana
<a href="http://cs.joensuu.fi/sipu/datasets/">http://cs.joensuu.fi/sipu/datasets/</a>	3	2	300	Path-Based1
<a href="http://cs.joensuu.fi/sipu/datasets/">http://cs.joensuu.fi/sipu/datasets/</a>	3	2	312	3-Spiral
<a href="http://sci2s.ugr.es/keel/category.php?cat=clas#sub2">http://sci2s.ugr.es/keel/category.php?cat=clas#sub2</a>	2	3	306	Haber man
<a href="http://sci2s.ugr.es/keel/category.php?cat=clas#sub2">http://sci2s.ugr.es/keel/category.php?cat=clas#sub2</a>	2	9	683	Wisconsin
<a href="http://sci2s.ugr.es/keel/category.php?cat=clas#sub2">http://sci2s.ugr.es/keel/category.php?cat=clas#sub2</a>	3	13	178	Wine

نتایج حاصل از شبیه‌سازی‌ها روی مجموعه داده‌های داده شده، در جدول ۲ آورده شده است. همان‌طور که مشاهده می‌شود الگوریتم پیشنهادی با سرعت بیشتر و حجم محاسبات کمتر توانسته به دقتی مشابه دقت الگوریتم رایج ALM دست یابد.

تاکنون در تمامی شبیه‌سازی‌های مرتبط با ALM، داده آموزشی، تا پایان فرآیند آموزش و همچنین در طی فرآیند مدلینگ ثابت باقی می‌ماند. اما برچسب زمانی (فاکتور زمان) روی داده‌های ورودی آموزشی حائز اهمیت بوده و داده‌هایی که دیرتر آموزش دیده‌اند نقش بیشتری در فرآیند تصمیم‌گیری نهایی خواهند داشت. افزودن این توانمندی به الگوریتم یادگیری می‌تواند برای کاربردهای غیرایستادن و متغیر با زمان مناسب باشد. بنابراین در روش پیشنهادی یک ضریب فراموشی برای صفحات IDS، که اثرات قطرات جوهر از صفحات IDS پاک شوند، قرار خواهیم داد. با افزودن ضریب فراموشی، این قابلیت به وجود خواهد آمد که در جاهایی که قطره بیشتر چکانده شده، فراموشی کمتر و یا مابه‌ازای پیاده‌سازی اسپایکی آن، ضریب سیناپسی قوی‌تر خواهد بود. قطره‌ای که چکانده می‌شود با گذشت زمان، ارتفاعش کاهش یافته و کم‌کم ناپدید می‌گردد و اگر قطره‌ای دیگر روی آن چکانده نشود، اطلاعات از بین خواهد رفت. این قابلیت معادل با قانون هب است که فعالیت هم‌زمان دو نورون (معادل با تکرار رخدادها) سبب تقویت ضریب سیناپسی بین نورون‌ها خواهد شد.

با فرض الگوی آموزشی  $z(t)$  در صفحه IDS، برای پیاده‌سازی با نگاهی مشابه ادبیات فیلتر وقتی RLS، صفحه IDS وابسته به پارامتر زمان به صورت زیر تعریف می‌کنیم:

$$IDS(z(t)) = \sum_{i=0}^t \lambda^{t-i} \times z(i) \quad (11)$$

$$IDS(z(t)) = \lambda^t \times z(0) + \lambda^{t-1} \times z(1) + \dots + \lambda^0 \times z(t) \quad (12)$$

در این رابطه،  $\lambda$  بیانگر ضریب فراموشی است. این رابطه در نگاه اول پیچیده به نظر می‌رسد و با افزایش زمان به حافظه زیادی احتیاج دارد. ولی با نگاه دقیق‌تر به این رابطه درمی‌یابیم که یک فرم بازگشتی برای این رابطه مشابه فیلتر RLS وجود دارد:

$$IDS(z(1)) = \lambda \times z(0) + z(1) \quad (13)$$

$$IDS(z(t)) = z(t) + \sum_{i=0}^{t-1} \lambda^{t-i-1} \times z(i) \quad (14)$$

با بهره‌گیری از رابطه ۱۱ در نهایت خواهیم داشت:

$$IDS(z(t)) = z(t) + \lambda \times IDS(z(t-1)) \quad (15)$$

بنابراین رابطه ۱۵ بیان می‌دارد که مقدار قطره جوهر یک صفحه IDS، برابر با مجموع مقادیر قطره جوهری که هم‌اکنون آمده با یک ضریب فراموشی از مقادیر قطره جوهر در لحظات قبل است. این رابطه بیانگر این است که در طول زمان اثر داده‌های قبل با ضریب فراموشی کاهش داده می‌شود. بنابراین اندازه شعاع پخش جوهر را می‌توان بیانی از تطبیق‌پذیری دانست، به علت اینکه از

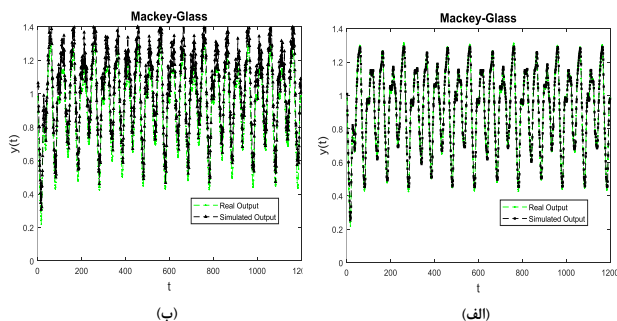


جدول ۲: نتایج حاصل از مقایسه مدل پیشنهادی با روش ANFIS و MLP روی انواع مجموعه داده‌ها.

MFs=number of membership function, Ep=epochs, Top=topology of network, Lr=learning rate.

Wine	Wisconsin	Haberman	3-Spiral	Path-Based1	Banana	الگوریتم
$r = 7$ $ACC = 98.34 \pm 2.31$	$r = 12$ $ACC = 96.76 \pm 1.57$	$r = 8$ $ACC = 74.12 \pm 6.97$	$r = 18$ $ACC = 99.64 \pm 0.43$	$r = 8$ $ACC = 98.83 \pm 1.75$	$r = 16$ $ACC = 89.86 \pm 1.86$	ALM
$MFs = genfis2/-$ $Ep = 70$ $ACC = 100 \pm 0.0$	$MFs = genfis2/-$ $Ep = 300$ $ACC = 89.87 \pm 0.63$	$MFs = genfis2/-$ $Ep = 200$ $ACC = 77.14 \pm 7.21$	$MFs = genfis3/-$ $Ep = 400$ $ACC = 99.65 \pm 0.65$	$MFs = genfis1/3$ $Ep = 350$ $ACC = 98.37 \pm 2.22$	$MFs = genfis1/3$ $Ep = 800$ $ACC = 73.21 \pm 2.16$	ANFIS
$Top = 13 - 22 - 3$ $Ep = 100$ $Lr = 0.05$ $ACC = 98.79 \pm 2.14$	$Top = 9 - 30 - 2$ $Ep = 200$ $Lr = 0.06$ $ACC = 94.57 \pm 1.34$	$Top = 3 - 4 - 2$ $Ep = 100$ $Lr = 0.03$ $ACC = 77.24 \pm 6.47$	$Top = 2 - 18 - 3$ $Ep = 500$ $Lr = 0.03$ $ACC = 97.85 \pm 4.88$	$Top = 2 - 27 - 3$ $Ep = 60$ $Lr = 0.075$ $ACC = 96.57 \pm 4.09$	$Top = 2 - 18 - 2$ $Ep = 200$ $Lr = 0.04$ $ACC = 74.21 \pm 2.79$	ALM
$\lambda = 0.998$ $r = 7$ $ACC = 98.86 \pm 2.47$	$\lambda = 1$ $r = 16$ $ACC = 96.76 \pm 1.56$	$\lambda = 0.999$ $r = 7$ $ACC = 73.47 \pm 6.23$	$\lambda = 0.997$ $1\theta r =$ $ACC = 99.45 \pm 0.57$	$\lambda = 0.998$ $r = 8$ $ACC = 98.84 \pm 1.76$	$\lambda = 0.99$ $r = 16$ $ACC = 89.98 \pm 1.72$	مدل پیشنهادی

موراکامی در مقالات [۲۱] و [۲۷] بیان کرده است که روش ALM با هر دو روش آموزش نمونه‌ای و مجموعه‌ای جواب‌های یکسانی دارد و بسته به نوع کاربرد می‌توان یکی از این دو مورد را استفاده قرار داد ولی سرعت روش مجموعه‌ای بیشتر است. برای شبیه‌سازی در روش ALM، مدل پیشنهادی، ANFIS و MLP نمونه-های آموزشی به صورت  $(y_{n-2}, y_{n-1}, y_n)$  انتخاب شده‌اند (یک سیستم دو-ورودی تک-خروجی). در جدول ۳، مقایسه روش‌ها آورده شده است. این شبیه‌سازی، ۱۰۰۰ داده به‌عنوان داده‌های آموزشی و ۲۰۰ داده به‌عنوان داده‌های تست بهره گرفته شده است و تعداد افزاها در دامنه ورودی‌ها ۶ است و خروجی ناشی از شبیه‌سازی در شکل ۶ آورده شده است.



شکل ۶ (الف) خروجی حاصل از الگوریتم پیشنهادی به ازای  $\lambda = 1$  (ب) خروجی حاصل از الگوریتم پیشنهادی به ازای  $\lambda = 0.99$

برای بررسی دقت الگوریتم و خطای مدل‌سازی از معیارهای FVU<sup>۴</sup> و همبستگی<sup>۵</sup> بیان شده در فرمول‌های ۱۷ و ۱۸ بهره گرفته شده است [۲۰] [۲۸]:

$$FVU = \frac{\sum_{i=1}^k (\hat{y}(x_i) - y(x_i))^2}{\sum_{i=1}^k (y - \bar{y})^2}, \bar{y} = \left(\frac{1}{k}\right) \sum_{i=1}^k y(x_i) \quad (17)$$

$$Corr = \frac{\sum_{i=1}^k (y_i - \bar{y}) \times (y_i^e - \bar{y}^e)}{\sqrt{\sum_{i=1}^k (y_i - \bar{y})^2 \times \sum_{i=1}^k (y_i^e - \bar{y}^e)^2}} \quad (18)$$

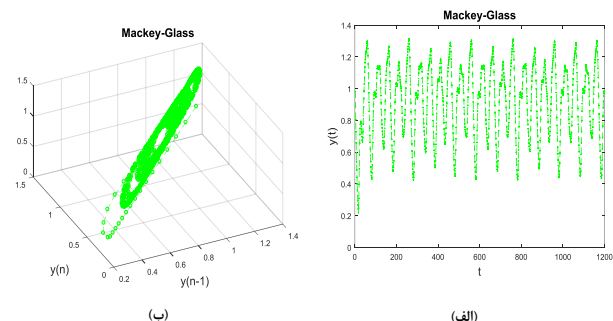
همان‌طور که از میزان انحراف معیار می‌توان نتیجه‌گیری کرد، این روش‌ها در مقایسه با روش شبکه‌های عصبی دارای پایداری بیشتر و نوسان کمتری می‌باشند. مکانیزم فراموشی، سبب می‌شود که در مکان‌هایی که بخش جوهر کمتری صورت می‌گیرد، اثرات داده‌ها از بین رفته و با همین مکانیزم اثرات مرتبط با نویزها حذف گردد.

### ۲-۵- مدل سازی

در این بخش، به عنوان اولین مثال، برای بررسی عملکرد روش پیشنهادی به بررسی آن روی دیتاست Mackey-Glass پرداخته شده است. این تابع سری زمانی، رفتاری شبه‌پریودیک با پریود مختلف در بازه‌های زمانی مختلف را دارا است (در این رابطه  $\tau=2$  انتخاب شده است که البته بسته به مقادیر آن شکل‌هایی با تفاوت ظاهری اندکی ایجاد می‌گردد).

$$\frac{dy}{dt} = \frac{2y(t-2)}{1+y(t-2)^{9.65}} - y(t) \quad (16)$$

در شکل ۵- الف و ب) نمایشی از سری Mackey-Glass به ازای ۱۲۰۰ نمونه زمانی از این تابع آورده شده است.

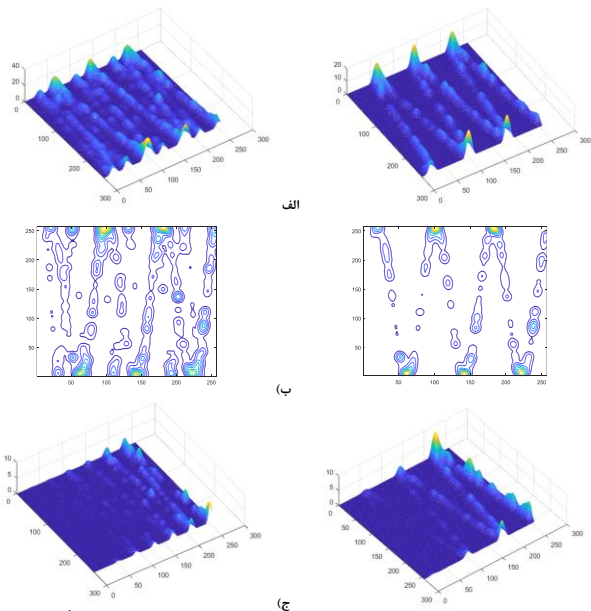


شکل ۵ (الف) نمودار تغییرات سری Mackey-Glass بر حسب زمان، (ب) نمودار دو بعدی سری زمانی Mackey-Glass.

در شبیه‌سازی، آموزش در هر دو روش (ALM رایج و مدل پیشنهادی) به صورت روش نمونه‌ای است. قابل‌بیان است که

آموزشی است که این امر مخرب است. در شکل ۷) به‌ازای دو حالت وجود ضریب فراموشی و عدم وجود ضریب فراموشی، نمایشی از پخش قطرات جوهر برای یک سیستم تک-ورودی تک-خروجی (تابع سینوسی که فرکانس آن با زمان افزایش می‌یابد) نمایش داده شده است.

در شکل ۷-الف) قابل مشاهده است که با آمدن نمونه‌های تابع سینوسی دوم که دارای فرکانس بیشتری نسبت به الگوی سینوسی اول است، شکل نامشخصی در صفحه IDS ایجاد شده است (فرکانس با گذشت زمان دو برابر شده است). در شکل ۷-ج) افزودن ضریب فراموشی، سبب شده که الگوی بهتری از تابع سینوسی باقی بماند. ضمناً افزایش بسیار زیاد ضریب فراموشی سبب می‌شود که الگوی ذخیره‌شده در صفحات IDS حذف شود و عملاً دقت کاهش یابد



شکل ۷) الف) صفحات IDS، تابع سینوسی با فرکانس متغیر با زمان (عدم بهره‌گیری از ضریب فراموشی)، ب) کانتورهای مرتبط با صفحات IDS، ج) بهره‌گیری از ضریب فراموشی در صفحات IDS.

با توجه به شبیه‌سازی‌های صورت گرفته، می‌توان نتیجه گرفت که افزودن مکانیزم فراموشی سبب می‌شود:

- ۱) روش پیشنهادی همانند الگوریتم اصلی دارای عملکرد مناسب است و صرفاً یک معیار فراموشی به الگوریتم افزوده شده است.
- ۲) در روش پیشنهادی، روزرسانی سیستم در هر زمان ممکن است. در صورت آمدن داده‌های جدید روزرسانی به‌صورت سراسری رخ نداده و به‌صورت محلی صورت می‌پذیرد.
- ۳) در روش پیشنهادی برخلاف روش‌هایی نظیر شبکه‌های عصبی با الگوریتم یادگیری پس انتشار خطا، دور از محاسبات پیچیده

خروجی حاصل از مدل  $\bar{y}$  و  $\bar{y}^e$  بیانگر میانگین بردارهاست و  $k$  بیانگر تعداد داده‌های تست است.  $FVU$  هر چه کمتر باشد بیانگر افزایش دقت مدل و معیار همبستگی هر چه به یک نزدیک تر باشد بیانگر دقت بیشتر خروجی خواهد بود.  $FVU$  رابطه مستقیم با  $MSE$  داشته و هر دو جهت حرکت یکسانی دارند (داشتن ارتباط مستقیم).

جدول ۳: نتایج حاصل از مقایسه مدل پیشنهادی با روش ALM رایج، ANFIS و MLP روی مجموعه داده Mackey-Glass.

الگوریتم	پارامترها	$FVU$	$PCC$	زمان (ثانیه)
ALM	$r = 6$	0.0447	0.9781	0.1823
ANFIS	$MFs = genfis1/2$ $Ep = 200$	0.0423	0.09771	0.3247
MLP	$Top = 2 - 10 - 1$ $Ep = 200$ $Lr = 0.09$	0.0449	0.9801	1.8927
مدل پیشنهادی	$\lambda = 0.99$ $r = 6$	0.0507	0.9762	0.1901

همان‌طور که در جدول ۳ مشاهده می‌شود، افزودن مکانیزم فراموشی در مجموعه داده فوق، هیچ مزیتی نسبت به روش ALM رایج ندارد و حتی اعمال فراموشی بر روی این مجموعه داده سبب کاهش دقت می‌شود. دلیل این اتفاق ناشی از عدم تغییر الگوی درونی داده‌ها است. در ادامه به بررسی روش پیشنهادی در فرآیند مدل‌سازی یک تابع سینوسی که فرکانس آن با زمان تغییر می‌کند، خواهیم پرداخت و نتایج مقایسه روش‌ها در جدول ۴، آورده شده است.

جدول ۴: نتایج حاصل از مقایسه ALM با مدل پیشنهادی، ANFIS و MLP در فرآیند مدل‌سازی یک تابع سینوسی.

الگوریتم	پارامترها	$FVU$	$PCC$
ALM	$r = 7$	1.1356	0.3309
ANFIS	$MFs = genfis1/2$ $Ep = 150$	0.0301	0.9851
MLP	$Top = 2 - 8 - 1$ $Ep = 200$ $Lr = 0.08$	0.0324	0.9832
مدل پیشنهادی	$\lambda = 0.99$ $r = 7$	0.0321	0.9847
مدل پیشنهادی	$\lambda = 0.96$ $r = 7$	0.0235	0.9894
مدل پیشنهادی	$\lambda = 0.81$ $r = 7$	0.2835	0.8499
مدل پیشنهادی	$\lambda = 0.22$ $r = 7$	0.8831	0.5645

طبق نتایج جدول، افزودن ضریب فراموشی، افزایش دقت را به همراه داشته است ولی همان‌طور که مشاهده می‌شود با کاهش بیشتر مقدار این ضریب، دقت به‌طور قابل‌توجهی کاهش یافته است و دلیل آن فراموش شدن سریع الگوی درونی داده‌های

Conference on Digital Manufacturing & Automation. 2010. IEEE.

[7]. French, R.M., Catastrophic forgetting in connectionist networks. Trends in cognitive sciences, 1999. 3(4): p. 128-135.

[8]. Averkin, A. and S. Yarushkevich, Review of research in the field of developing methods to extract rules from artificial neural networks. Journal of Computer and Systems Sciences International, 2021. 60(6): p. 966-980.

[9]. Nakayama, H. and K. Yoshii. Active forgetting in machine learning and its application to financial problems. in Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium. 2000. IEEE.

[10]. Salganicoff, M., Explicit Forgetting Algorithms for Memory Based Learning. 1993.

[11]. Panda, P., et al., Asp: Learning to forget with adaptive synaptic plasticity in spiking neural networks. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 2017. 8(1): p. 51-64.

[12]. Auge, D., et al., A survey of encoding techniques for signal processing in spiking neural networks. Neural Processing Letters, 2021. 53(6): p. 4693-4710.

[13]. Kato, A. and K. Morita, Forgetting in reinforcement learning links sustained dopamine signals to motivation. PLoS computational biology, 2016. 12(10): p. e1005145.

[14]. Shouraki, S.B. and N. Honda. Fuzzy controller design by an active learning method. in 31th symposium of intelligent control. Tokyo, Japan. 1998.

[15]. Merrikh-Bayat, F., S.B. Shouraki, and A. Rohani, Memristor crossbar-based hardware implementation of the IDS method. IEEE Transactions on Fuzzy Systems, 2011. 19(6): p. 1083-1096.

[16]. Afrakoti, I.E.P., S.B. Shouraki, and B. Haghghat, An optimal hardware implementation for active learning method based on memristor crossbar structures. IEEE Systems Journal, 2014. 8(4): p. 1190-1199.

[17]. Shahdi, S.A. and S.B. Shouraki. Supervised active learning method as an intelligent linguistic controller and its hardware implementation. in 2nd IASTED International Conference on Artificial Intelligence and Applications (AIA'02), Malaga, Spain. 2002.

[18]. Klidbary, S.H., et al. Outlier robust fuzzy active learning method (ALM). in 2017 7th international conference on computer and knowledge engineering (ICCKE). 2017. IEEE.

[19]. Klidbary, S.H., S.B. Shouraki, and I.E.P. Afrakoti, An adaptive efficient memristive ink drop spread (IDS) computing system. Neural Computing and Applications, 2019. 31(11): p. 7733-7754.

[20]. Jokar, E., et al., Hardware-Algorithm Co-Design of a Compressed Fuzzy Active Learning Method. IEEE Transactions on Circuits and Systems I: Regular Papers, 2020. 67(12): p. 4932-4945.

[21]. Klidbary, S.H. and S.B. Shouraki, A novel adaptive learning algorithm for low-dimensional feature space using memristor-crossbar implementation and on-chip training. Applied Intelligence, 2018. 48(11): p. 4174-4191.

[22]. Javadian, M., et al., Refining membership degrees obtained from fuzzy C-means by re-fuzzification. Iranian Journal of Fuzzy Systems, 2020. 17(4): p. 85-104.

[23]. سجاد حق زاد کلیدبری، سعید باقری شورکی، ارائه اپراتور جدید جایگزین پخش قطره جوهر در روش یادگیری فعال. مجله مهندسی برق دانشگاه تبریز، ۲۰۱۹.

[24]. Klidbary, S.H., S.B. Shouraki, and B. Linares-Barranco, Digital hardware realization of a novel adaptive ink drop spread operator and its application in modeling and classification and on-chip training. International Journal of Machine Learning and Cybernetics, 2019. 10(9): p. 2541-2561.

[25]. Javadian, M., A. Hejazi, and S.H. Klidbary, Obtaining Fuzzy Membership Function of Clusters With the Memristor Hardware Implementation and On-Chip Learning. IEEE Transactions on Emerging Topics in Computational Intelligence, 2022.

[26]. Murakami, M. and N. Honda, A study on the modeling ability of the IDS method: A soft computing technique using pattern-

است و نیازمند بهینه‌کردن یک تابع پیچیده نیست و هزینه پیاده‌سازی کمتری نسبت به سایر الگوریتم‌های محاسباتی موجود دارد.

## ۶- نتیجه‌گیری

در پیاده‌سازی عملی، به دلیل محدودیت فضا و تعداد نورون‌های شبکه، تعداد محدودی الگوی آموزشی، قابل آموزش به شبکه است و الگوهای بعدی با این الگوها تداخل مخرب پیدا خواهند کرد و همین امر نیازمندی به ارائه الگوریتمی مبتنی بر فراموشی را دوچندان می‌نماید. روش یادگیری فعال، در کاربردهای مختلف، دقت قابل قبولی را نسبت به سایر الگوریتم‌های یادگیری داشته است. بنابراین، مکانیزم فراموشی به صفحات IDS روش یادگیری فعال افزوده شده است تا الگوریتم مشابهت بیشتری با بیولوژیک بیابد. تاکنون در تمام پیاده‌سازی‌های روش یادگیری فعال، قطره جوهر بر روی صفحات پخش قطرات جوهر ثابت مانده و هیچ‌گونه فراموشی رخ نمی‌دهد، که این امر مغایر با بیولوژیک بوده و در این مقاله با افزودن ضریب فراموشی به صفحات IDS، قدرت محاسباتی روش پیشنهادی در برخورد با مجموعه داده‌های متغیر با زمان دوچندان شده است. می‌توان بیان داشت که با استفاده از مکانیزم افزوده شده (با کمترین پیچیدگی محاسباتی)، توانستیم خاصیت وفقی بودن ایجاد نماییم و همین امر افزایش کارایی الگوریتم را در پی داشته است. شبیه‌سازی‌های صورت گرفته، نشان از توانایی بالای این الگوریتم پیشنهادی دارد. در کارهای آینده، برای افزایش شباهت بیولوژیکی، باورپذیری، واقعیت‌پذیری و تقویت جنبه‌های ساختاری سیستم ارائه شده، فرآیند یادگیری را بر مبنای مدل‌های عصبی اسپایکی ارائه خواهد شد. امیدواریم با تلفیق روش‌های فازی و سیستم‌های عصبی به نتایج بهتری دست یابیم. ضمن این-که، روش ALM روی بسترهای مختلف سخت‌افزاری پیاده‌سازی شده است و افزودن مکانیزم فراموشی نیز روی این سخت‌افزارها، به ساده‌ترین صورت ممکن امکان‌پذیر خواهد بود.

## مراجع

[1]. Liu, F., et al., A New Fuzzy Spiking Neural Network Based on Neuronal Contribution Degree. IEEE Transactions on Fuzzy Systems, 2021.

[2]. Qasem, S.N. and A. Mohammadzadeh, A deep learned type-2 fuzzy neural network: Singular value decomposition approach. Applied Soft Computing, 2021. 105: p. 107244.

[3]. Hebb, D.O., The organization of behavior: A neuropsychological theory. 2005: Psychology Press.

[4]. Fox, M.D. and M.E. Raichle, Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. Nature reviews neuroscience, 2007. 8(9): p. 700-711.

[5]. Abbott, L.F. and S.B. Nelson, Synaptic plasticity: taming the beast. Nature neuroscience, 2000. 3(11): p. 1178-1183.

[6]. Pan, H.-j., et al. A Category Theory Model for Learning and Memory of the Human Brain. in 2010 International

- based information processing. International journal of approximate reasoning, 2007. 45(3): p. 470-487.
- [27]. Murakami, M. and N. Honda. Classification performance of the IDS method based on the two-spiral benchmark. in 2005 IEEE International Conference on Systems, Man and Cybernetics. 2005. IEEE.
- [28]. Hwang, J.-N., et al., Regression modeling in back-propagation and projection pursuit learning. IEEE Transactions on neural networks, 1994. 5(3): p. ۳۴۲-۳۵۳ .

## پاورقی‌ها:

---

- <sup>1</sup> *Sensory Memory*
- <sup>2</sup> *Short-Term Memory*
- <sup>3</sup> *Long-Term Memory*
- <sup>4</sup> *Memory-Based Learning (MBL)*
- <sup>5</sup> *Narrow Path (NP)*
- <sup>6</sup> *Spread (SP)*
- <sup>7</sup> *Recursive Least Square (RLS)*
- <sup>8</sup> *Fraction of Variance Unexplained (FVU)*
- <sup>9</sup> *Pearson Correlation Coefficients (PCC)*
- <sup>10</sup> *Mean Square Error (MSE)*
- <sup>11</sup> *Error Back Propagation (EBP)*