

# کنترل آونگ وارون با بهره‌گیری از داده‌های منفی

سجاد حق زاد کلیدبری<sup>۱</sup>

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

[s.haghzad@znu.ac.ir](mailto:s.haghzad@znu.ac.ir)

**چکیده:** در فرآیند آموزش الگوریتم‌های یادگیری، همواره داشتن مجموعه داده‌های آموزشی مناسب مورد توجه است. وجود داده‌های پرت، داده‌های نویزی و داده‌های نامناسب، همواره بر عملکرد الگوریتم‌های موجود تأثیرگذار است. روش یادگیری فعال<sup>۱</sup>، یکی از ابزارهای قدرتمند در محاسبات نرم است که از نحوه عملکرد مغز انسان بهره می‌گیرد. عملکرد این الگوریتم کاملاً مبتنی بر محاسبات ساده است و یک سیستم چند-ورودی تک-خروجی را به صورت مجموعی از سیستم‌های تک-ورودی تک-خروجی در نظر می‌گیرد و با شکستن مسئله پیچیده به چندین مسئله ساده‌تر موجب درک‌پذیری ساده آن می‌شود. در این مقاله برای بهبود عملکرد این الگوریتم، با تغییراتی در فرآیند آموزش، از داده‌هایی تحت عنوان داده‌های منفی استفاده شده است. در روش یادگیری فعال به واسطه استفاده از اپراتور پخش قطره جوهر<sup>۲</sup>، بهره‌گیری از داده‌های منفی در فرآیند آموزش امکان‌پذیر است. با این اقدام، سیاستی مشابه سیاست پاداش و جزاء در روش‌های یادگیری تقویتی<sup>۳</sup> اجراء می‌شود. شبیه‌سازی‌های انجام شده بیان از عملکرد مناسب و افزایش قدرت محاسباتی روش پیشنهادی نسبت به روش موجود دارد و مقدار FVU، ۰.۱۴۳، حاصل شده است. همچنین با روش پیشنهادی، افزایش دقت و سرعت در فرآیند کنترل سیستم پاندول معکوس<sup>۴</sup>، که یک مسئله کلاسیک در سیستم‌های کنترل می‌باشد، کاملاً مشهود است.

**واژه‌های کلیدی:** روش یادگیری فعال، اپراتور پخش قطره جوهر، سیستم استنتاج فازی، داده منفی، کنترل

نوع مقاله: پژوهشی

DOI: 10.52547/jiaeee.20.1.143

تاریخ ارسال مقاله: ۱۴۰۰/۱۰/۲۳

تاریخ پذیرش مشروط مقاله: ۱۴۰۱/۰۴/۲۹

تاریخ پذیرش مقاله: ۱۴۰۱/۰۵/۰۸

نام نویسنده‌ی مسئول: سجاد حق زاد کلیدبری

نشانی نویسنده‌ی مسئول: ایران - زنجان - کلیومتر ۶ جاده تبریز - دانشگاه زنجان - دانشکده مهندسی

## ۱- مقدمه

تاکنون الگوریتم‌های یادگیری بسیاری برای شبیه‌سازی عملکرد مغز انسان معرفی شده‌اند. الگوریتم‌های بیان شده در حل بسیاری از مسائل مانند تخمین توابع، پیش‌بینی، داده‌کاوی و مسائلی نظیر کاربرد دارند. بهره‌گیری از تکنیک فازی، همواره مورد توجه بوده و از مهم‌ترین مزیت‌های روش‌های حل مسئله فازی نسبت به روش‌های کلاسیک بهره‌گیری از متغیرهای زبانی و رابطه‌های "اگر و آنگاه" و عدم استفاده از روابط پیچیده است. در این روش‌ها سعی می‌شود که مسئله، مانند انسان با استفاده از زبان طبیعی مطرح و حل شود.

سیستم پاندول معکوس یکی از مسائل کنترلی بسیار مطرح بوده و از سال ۱۹۵۰، پاندول معکوس برای تدریس تئوری کنترل فیدبک مورد استفاده قرار گرفته شده است. اولین حل این مسئله در سال ۱۹۶۰ بیان شده است و تقریباً بیش از ۵۰ سال است که در تئوری کنترل مورد بررسی قرار می‌گیرد [۱]. سیستم پاندول معکوس دارای ویژگی چند متغیره، غیرخطی و ناپایدار است که یک سیستم دینامیکی ناپایدار مطلق است. این سیستم در واقع یک سیستم تک-ورودی دو-خروجی بوده که باید بتوان با یک سیگنال کنترلی که به موتور ارائه داده می‌شود، موقعیت ارا به و زاویه پاندول را تحت کنترل داشته باشیم. این سیستم دارای دو نقطه تعادل پایدار و ناپایدار می‌باشد. تاکنون مقالات زیادی برای بررسی پایداری و کنترل پاندول معکوس ارائه شده است [۲-۷] و رفرنس‌های موجود در این مقالات.

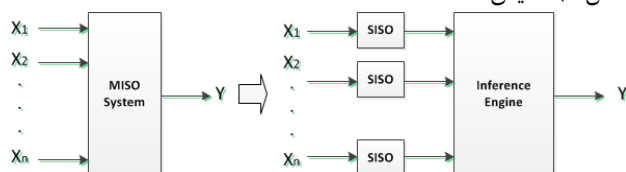
یکی از مهم‌ترین روش‌ها در کنترل پاندول معکوس، بهره‌گیری از تکنیک فازی می‌باشد. هدف در کنترل فازی، مدل‌سازی فرد خبره است [۸، ۹]. دانش این فرد خبره را نمی‌توان توسط سیستم‌های کلاسیک مدل نمود و سیستم‌های فازی جهت مدل‌سازی اپراتور انسانی مناسب‌تر است. به دلیل اینکه دانش این کار را فرد خبره، بر اساس تجربه کسب نموده است و به راحتی توسط معادلات تفاضلی یا دیفرانسیلی بیان نمی‌گردد. از مشکلات متداول در هدایت پاندول این است که بتواند از شرایط اولیه به حالت تعادل برسد و ارا به را در موقعیت مطلوب ثابت شود. هدف از این مقاله این است که با بهره گرفتن از داده منفی، وضعیت پاندول را از حالت نامتعادل به نقطه تعادل برساند. بنابراین برای داشتن کنترلی مشابه خبرگی انسانی، از الگوریتم‌هایی با عملکرد مشابه مغز انسان می‌توان بهره گرفت. به نظر می‌رسد که انسان در برخورد با یک سیستم چند-ورودی تک-خروجی، آن را به صورت مجموعه‌ای از چند سیستم تک-ورودی تک-خروجی در نظر می‌گیرد. یکی از روش‌هایی که بر این مفاهیم استوار است، روش ALM نام دارد و در سال ۱۹۹۷ ارائه شده است [۱۰]. در این الگوریتم از اپراتور IDS برای مدل‌سازی عدم قطعیت در داده‌ها بهره گرفته می‌شود که محاسبات بسیار ساده‌ای را دارد. در این اپراتور پخش قطرات جوهر با شعاع مناسب به واسطه وجود هر داده آموزشی صورت می‌گیرد، و با این اقدام مدل‌سازی فضای تجربه انجام می‌پذیرد.

مراحل طراحی کنترلر فازی از قسمت‌های، فازی‌سازی ورودی‌ها و تعریف توابع تعلق، تعریف قوانین استنتاج، غیرفازی‌سازی خروجی تشکیل شده است. با استفاده از الگوریتم یادگیری فعال، قصد داریم یک کنترلر فازی برای کنترل پاندول معکوس ارائه دهیم. در این الگوریتم بدون بهره‌گیری از روابط ریاضی (بدون نیاز به دانستن قوانین مکانیکی حاکم) و صرفاً با بکارگیری هم زمان داده مثبت (قطره جوهر معمولی) و داده منفی (قطره جوهر با عملکرد عکس) به مدل‌سازی فضای تجربه می‌پردازیم. داده منفی سبب می‌شود که پخشی نامتقارن در فضای IDS داشته باشیم، یعنی در جهتی که داده مناسب مشاهده شده، شعاع پخش‌شدگی جوهر بیشتر و در جهتی که داده مناسب مشاهده نشده، شعاع پخش‌شدگی جوهر کمتر باشد. بنابراین با این اقدام، سیاستی مشابه سیاست پاداش و جزاء در روش‌های یادگیری تقویتی را خواهیم داشت که افزایش قدرت محاسباتی را در پی داشته و افزایش دقت و سرعت آن در فرآیند کنترل پاندول معکوس کاملاً مشهود خواهد بود.

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

## ۲- روش یادگیری فعال (ALM)

نحوه انجام محاسبات در مغز انسان ماهیتی کیفی و نادقیق داشته و همین امر نیازمندی به روش محاسباتی نرم با بهره‌گیری از عدم قطعیت را در شبیه‌سازی رفتار مغز دوچندان می‌کند. عدم قطعیت، جزء جدانشدنی پدیده‌های اطراف ما قلمداد شده و سیستم‌های فازی توانسته‌اند تا حد بسیار مناسبی در مدل‌سازی عدم قطعیت عملکرد مناسبی داشته باشند. روش یادگیری فعال از جمله الگوریتم‌های کارا و مؤثر در حوزه فازی است. این الگوریتم و توسعه‌یافته‌های آن در حوزه‌های مختلفی نظیر کنترل [۱۱، ۱۲]، رباتیک [۱۳]، مدل‌سازی [۱۴-۱۷]، دسته‌بندی [۱۸-۲۰]، خوشه‌بندی [۲۱، ۲۲] استفاده شده است. عملکرد ALM برای یک سیستم چند-ورودی تک-خروجی و شکستن مسئله به چند زیر سیستم تک-ورودی تک-خروجی در شکل (۱) نمایش داده شده است.



شکل (۱): عملکرد روش یادگیری فعال در برخورد با سیستم‌های

چند-ورودی تک-خروجی

در شکل ۳) آورده شده، الگویی پس از اعمال اپراتور پخش قطره جوهر و گسترانده شدن قطرات جوهر بر روی آن ایجاد شده است. از هر یک از صفحات IDS دو ویژگی مسیر باریک<sup>۵</sup> و شعاع پراکندگی<sup>۶</sup> استخراج می‌شود. در الگوی ایجاد شده، به‌ازای یک نقطه تست دلخواه، مقدار مسیر با رنگ قرمز و مقدار شعاع پراکندگی نمایش داده شده است.

## ۲-۲- بخش استنتاج

همانطور که بیان شد، بعد از اعمال اپراتور پخش قطره جوهر بر روی صفحات دو بعدی، با استفاده از بخش استخراج ویژگی، ویژگی‌های مسیر باریک و شعاع پراکندگی از الگوهای ایجاد شده بر روی صفحات IDS استخراج می‌گردند که در مرحله استنتاج الگوریتم ALM مورد استفاده قرار می‌گیرند. مسیر باریک بیانگر رابطه بین ورودی و خروجی سیستم است و شعاع پراکندگی بیانگر تأثیرگذاری متغیر ورودی  $x$  در مقایسه با سایر متغیرهای ورودی، در تعیین مقدار خروجی است. روش‌های مختلفی برای محاسبه مسیر باریک وجود دارد که می‌توان به استفاده از عملگر ماکزیمم، عملگر میانگین، مقایسه مقادیر مجموع متوالی شدت جوهر اشاره نمود. برای یافتن شعاع پراکندگی حول هر نقطه مسیر باریک می‌توان غلظت جوهر هر نقطه یا شعاع پخش شدگی حول هر نقطه را در نظر گرفت. با فرض مجموعه  $S$  که شامل  $N$  داده آموزشی که ابعاد آن‌ها  $D$  است، می‌توان از روابط بیان‌شده بهره گرفت.

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

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

اولین گام در روش ALM، کوانتیزه نمودن متغیرهای ورودی و خروجی است (برای شبیه‌سازی سریع و پیاده‌سازی سخت‌افزاری آسان‌تر). با فرض سطوح کوانتیزاسیون به ترتیب برابر با  $R_{max}$  و  $R_{min}$ ، که معمولاً در شبیه‌سازی‌ها این مقادیر برابر در نظر گرفته می‌شود، داریم:

$$F_i(X_{iq}) = \left\lfloor \frac{(x_i - x_{min}^i)}{\Delta x^i} \right\rfloor + 1, \quad x_i \in [x_{min}^i, x_{max}^i], \quad \Delta x^i = \frac{x_{max}^i - x_{min}^i}{R_{max}} \quad (3)$$

$$H(Y_q) = \left\lfloor \frac{(y - y_{min})}{\Delta y} \right\rfloor + 1, \quad y \in [y_{min}, y_{max}], \quad \Delta y = \frac{y_{max} - y_{min}}{R_{min}} \quad (4)$$

در نهایت، مقادیر کوانتیزه شده بصورت زیر خواهند بود:

$$Eq. (3) \quad (5)$$

$$(x_{i1}, x_{i2}, \dots, x_{iD})^T \rightarrow X_{iq} \in \{1, 2, \dots, R_{max}\},$$

$$Eq. (4) \quad (6)$$

$$(y_1, y_2, \dots, y_D)^T \rightarrow Y_q \in \{1, 2, \dots, R_{min}\},$$

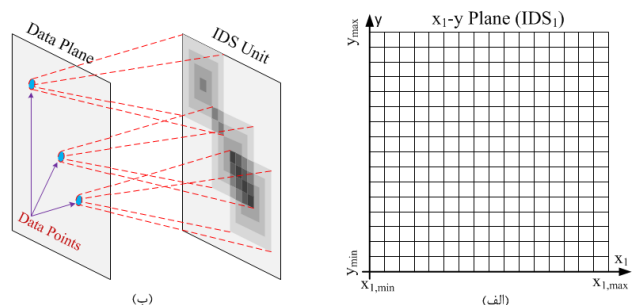
با فرض  $p(x, y)$  یک نقطه در صفحه IDS باشد و تیرگی در این نقطه از صفحه با  $d(x, y)$  بیان گردد و قطره جوهر همانند یک تابع گاوسی با واریانس مناسب باشد که در مرکز این تابع گاوسی مقدار برابر با یک است. در این صورت بروزرسانی صفحه IDS پس از اعمال

روش یادگیری فعال دارای دو بخش اساسی، ایجاد صفحات IDS و بخش استنتاج می‌باشد که در بخش‌های بعدی توضیح داده خواهد شد.

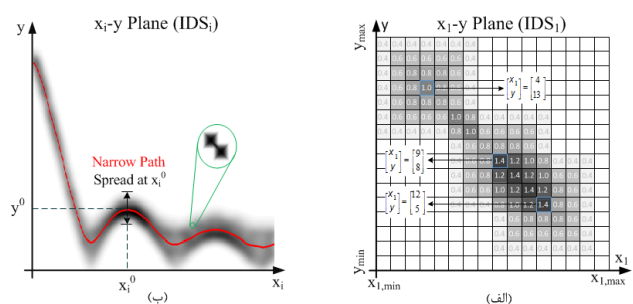
## ۲-۱- اپراتور IDS

موتور اصلی روش یادگیری فعال اپراتور پخش قطره جوهر است که عدم قطعیت را مدل‌سازی می‌کند. 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 و رفتاری سیستم سعی در تقلید و شبیه‌سازی رفتار مغز دارد. این عدم اطمینان و نگاه نادقیق به رخداد‌های اطراف با استفاده از یک عملگر تحت نام اپراتور IDS بیان می‌شود. عملکرد اپراتور IDS از این واقعیت نشأت می‌گیرد که تجربیات ماهیتی پیوسته دارند. خواص یک نقطه مشاهده شده منحصراً متعلق به همان نقطه نبوده و همسایگی آن نقطه را نیز شامل می‌شود و با فاصله‌گیری از نقطه تجربه، درجه اطمینان ما به هم خواص نقاط اطراف تجربه کاهش می‌یابد.

در روش ALM بعد از شکستن سیستم به زیرسیستم‌ها، هر کدام از زیر سیستم‌های تک- ورودی تک- خروجی 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. در شکل ۲) نمایشی از عملکرد اپراتور IDS آورده شده است.

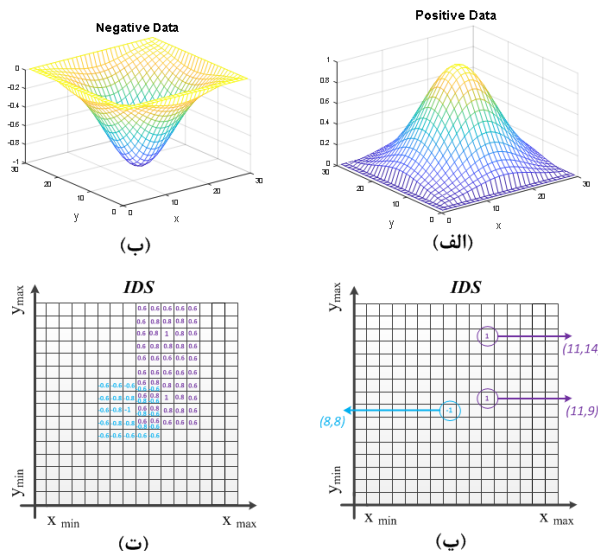


شکل (۲): (الف) صفحه مشبک خالی از قطره جوهر (ب) داده‌های آموزشی

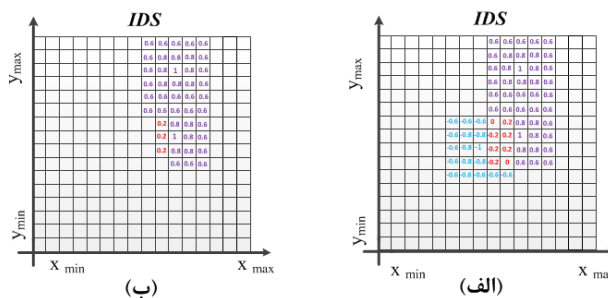


شکل (۳): (الف) اعمال اپراتور IDS روی صفحات IDS (ب) واحد استخراج کننده ویژگی [۱۷]

گردد. فرآیند گسترش یک داده منفی همانند داده مثبت است و فقط با این تفاوت که گسترش یک داده منفی نیز، منفی خواهد بود. برای تشریح بهتر عملکرد داده منفی، مثالی در شکل ۴ آورده شده است. همان‌طور که در شکل ۴ مشاهده می‌شود ابتدا قطره‌های جوهر مثبت و سپس قطره منفی روی صفحات پخش شده و سپس جمع جبری قطره‌های در شکل ۵) با اعمال جمع جبری، مقادیر منفی باقیمانده در صفحه IDS را به علت اینکه از جنس درجه تعلق هستند و باید مقادیری بین ۰ و ۱ داشته باشند روی مقادیر ۰ ثابت می‌کنیم (مقادیر صفر در سلول‌های IDS بیانگر این هستند که صفحات IDS خالی از قطره جوهر هستند و هیچ‌گونه پخشی در این سلول‌ها رخ نداده است).



شکل ۴: (الف) پخش قطره جوهر مثبت، (ب) پخش قطره جوهر منفی، (پ) داده مثبت و داده منفی، (ت) پخش توأمان قطره جوهر مثبت و منفی



شکل ۵: (الف) اعمال جمع جبری در صفحه IDS، (ب) ایجاد پخشی نامتقارن و الگوی متناسب ایجاد شده

همان‌طور که مشاهده می‌شود، داده منفی سبب ایجاد ناهمگنی شده است و این ناهمگنی در فضای IDS یعنی اینکه یک سمت‌هایی پخش جوهر راحت‌تر و در سمت‌های دیگر سخت‌تر صورت می‌پذیرد. در این بخش قصد داریم، اثرات داده منفی را در پخش قطرات در نظر گرفته و دقت آن‌را در فرآیند کنترل پاندول معکوس بررسی کنیم.

داده آموزشی جدید در نقطه  $(x_s, y_s)$  به صورت زیر قابل بیان خواهد بود.

$$p_{x,y} = \{p(x, y) \mid x \in X_i, y \in Y\} \quad (7)$$

$$d(x_s + u, y_s + v) = d(x_s + u, y_s + v) + h(u, v), \quad (8)$$

$$-R < u, v < R$$

در رابطه فوق R بیانگر شعاع قطره جوهر و h شکل تابع قطره جوهر است. با استفاده از تعاریف بیان شده می‌توان توابع  $\psi$  و S را به ترتیب برای یافتن مسیر باریک و شعاع پراکندگی در صفحه IDS تعریف نمود:

$$\psi_{x_i}(x) = \{b \mid \sum_{y=y_{\min}}^b d(x, y) \approx \sum_{y=b}^{y_{\max}} d(x, y), b \in Y\} \quad (9)$$

$$\sigma_{x_i}(x) = \max_{y \in Y} \{y \mid d(x, y) > T\} - \min_{y \in Y} \{y \mid d(x, y) > T\} \quad (10)$$

رابطه اول بیانگر مسیر باریک، به ازای هر ورودی x برابر با b است به شرطی که در صفحه دو بعدی، مجموع غلظت جوهرهای ذخیره شده در سلول‌های بالای مختصات  $(x, b)$  برابر با مجموع غلظت جوهرهای ذخیره شده در سلول‌های پایین مختصات  $(x, b)$  باشد. رابطه دوم مقدار تابع پراکندگی در نقطه‌ای مانند x متناسب با عرض مؤثر الگوی قرار گرفته (با تعداد سلول‌های دارای غلظت قطره جوهر) بر روی صفحه IDS در ستون متناظر با x حول مسیر باریک است. پارامتر T بیانگر مقدار آستانه برای یافتن و تأثیرگذاری مقدار تیرگی ذخیره شده در سلول برای مقدار پراکندگی است. در نهایت روش ALM بر مبنای مقادیر فوق برای یافتن مقدار خروجی جمع‌وزن‌داری را بدست می‌آورد.

بنابراین نحوه عملکرد اپراتور قطره جوهر عامل بسیار مهمی در عملکرد، کارایی، خطا خروجی و سرعت همگرایی الگوریتم ALM است. یکی از مشکلاتی که روش IDS این است که داده‌ها را فقط به صورت متقارن پخش می‌کند. در این مقاله اثرات داده‌های منفی را در پخش قطرات جوهر برای ایجاد پخشی نامتقارن در نظر می‌گیریم. این راهکار همانند سیاست پاداش و جزاء در یادگیری تقویتی است و با افزودن این قابلیت به صفحات IDS و روش یادگیری فعال، دقت آن را در فرآیند کنترل پاندول معکوس بررسی می‌کنیم.

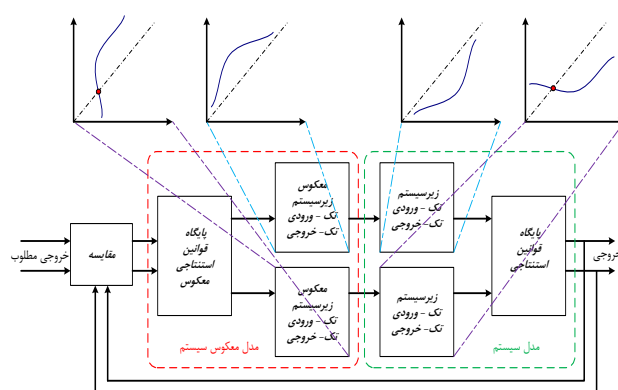
### ۳- الگوریتم پیشنهادی

در این بخش قصد داریم قصد داریم با بکارگیری هم‌زمان داده مثبت (قطره جوهر معمولی) و داده منفی (قطره جوهر با عملکرد عکس) به مدل‌سازی فضای تجربه پرداخته و بررسی کنیم که آیا این اقدام سبب افزایش یا کاهش قدرت محاسباتی می‌شود. داده منفی، به داده‌های نسبت داده می‌شود که تأثیر منفی دارند. داده منفی سبب می‌شود که پخشی نامتقارن در فضای IDS داشته باشیم، یعنی در جهتی که داده مناسب مشاهده شده، شعاع بیشتر و در جهاتی که داده مناسب مشاهده نشده، شعاع کمتر باشد. بنابراین قطره جوهر منفی برای پاک کردن شواهدی که از مقدار قبلی مدل پشتیبانی می‌کنند استفاده می‌شود. می‌توان بیان داشت که با این اقدام، داده با رزولوشن بیشتر یا همان شعاع کمتر در سمتی که تجربه مشاهده شده است، پخش می‌شود.

## ۴- پاندول معکوس

از جمله مهم‌ترین مزیت‌های روش‌های فازی نسبت به روش‌های کلاسیک، عدم استفاده از روابط پیچیده ریاضی است. در روش‌های طراحی کنترلر کلاسیک معمولاً سیستم به صورت یک معادله دیفرانسیل بیان می‌شود. در تئوری کنترل (کنترل کلاسیک)، فرآیند سنتز کنترلر به طور مستقیم امکان‌پذیر نیست، چرا که سیستم‌های چند- ورودی تک- خروجی عملاً معکوس پذیر نیستند. در صورتی که اگر چنین کاری امکان‌پذیر باشد، طراحی کنترلر بسیار راحت خواهد بود و این امکان در روش یادگیری فعال وجود دارد.

روش یادگیری فعال و با مدل‌سازی هر یک از زیرسیستم‌ها توسط صفحات IDS، هر یک از زیرسیستم را می‌توانیم معکوس کنیم و تأثیر آن را روی خروجی سیستم مشاهده کنیم (بنابراین در روش یادگیری فعال امکان معکوس‌پذیری مدل سیستم در مدل‌سازی وجود دارد و در این حالت ورودی‌ها نقش خروجی را بازی می‌کنند). شمای کلی طراحی کنترلر مبتنی بر الگوریتم ALM در شکل (۶) آورده شده است. با توجه به خطای خروجی و نقطه کار، مقادیر مناسب را برای یک سیستم فراهم می‌آورد تا در نهایت، خروجی نقطه هدف را دنبال نماید.

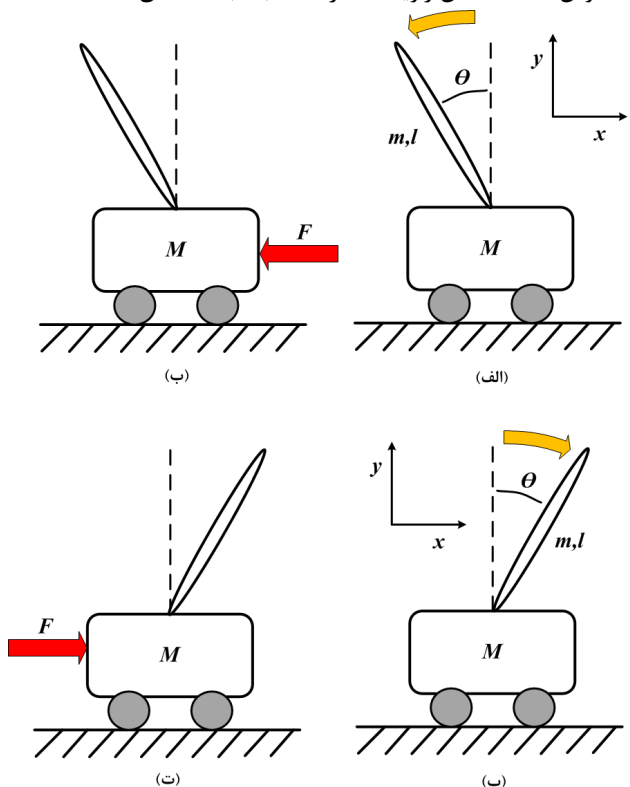


شکل (۶): شمای کلی طراحی کنترلر با بهره‌گیری از ALM

حفظ تعادل پاندول معکوس سوار روی یک ارابه متحرک که در راستای افقی حرکت می‌کند، یک مسئله کلاسیک و معروف در سیستم‌های کنترلی محسوب می‌شود. سیستم پاندول معکوس دو نقطه تعادل ذاتی دارد که یکی پایدار و دیگری ناپایدار است. نقطه تعادل پایدار، جایی است که پاندول روبه پایین قرار گرفته است و بدون نیاز به هیچ کنترل‌کننده‌ای، سیستم به طور طبیعی در این حالت قرار خواهد داشت. نقطه تعادل ناپایدار مربوط به وضعیتی می‌شود که پاندول دقیقاً روبه بالا قرار گرفته و در نتیجه نیاز به یک کنترلر برای حفظ تعادل دارد. هدف اصلی حفظ تعادل در نقطه ناپایدار است و هدف کنترلی این است که وضعیت پاندول از یک حالت نامتعادل به نقطه تعادل برسد و این وضعیت حفظ شود. نمای کلی از پاندول معکوس تک لینکه در شکل (۷) زیر آورده شده است.

طبق شکل (۷) و نحوه حرکت پاندول، باید بتوان با یک سیگنال کنترلی که به موتور داده می‌شود، به طور هم‌زمان موقعیت ارابه و زاویه پاندول را تحت کنترل داشت.

برای اینکه این پاندول، همانند شکل تعادل خود را حفظ کند، لازم است تا هنگامی که پاندول به سمت راست منحرف می‌شود، ارابه به سمت راست حرکت کند و بالعکس. منظور از کلمه تعادل در "پاندول معکوس"، نگه داشتن زاویه ۹۰ درجه نسبت به خط افق است.



شکل (۷): (الف) پاندول معکوس تک لینکه در حال سقوط در جهت نمایش داده شده، (ب) جهت حرکت ارابه برای ایجاد تعادل در پاندول، (پ) پاندول معکوس تک لینکه در حال سقوط در جهت نمایش داده شده و (ت) جهت حرکت ارابه برای ایجاد تعادل در پاندول

بعد از جمع‌آوری داده، سیستم تحت کنترل با استفاده از ALM مدل‌سازی می‌شود و منحنی رفتاری آن توسط صفحات IDS ذخیره و بیان می‌شود (فاز مدل‌سازی). سپس در فاز کنترل، خروجی‌های سیستم تحت کنترلر به عنوان ورودی به کنترلر اعمال شده و از روی صفحات IDS خروجی کنترلر محاسبه شده و به سیستم اصلی اعمال می‌شود. یعنی مؤثرترین ورودی‌ها با کمترین پراکندگی در صفحه IDS یافت می‌شود و بر مبنای آن‌ها صفحات IDS و قوانین فازی روش ALM بدست می‌آید. در این روش، در ابتدای امر یک ارزیابی روی داده‌های جمع‌آوری شده صورت می‌پذیرد و این داده‌ها به دو دسته داده‌های مناسب (خروجی مطلوب) و داده‌های غیرمناسب (خروجی نامطلوب) تفکیک می‌شوند. داده مثبت نمایانگر حرکت مناسب و داده منفی نمایانگر حرکت غیرمناسب است. حرکت مناسب حرکتی است که منجر به انحرافی کمتر از مقدار اولیه در خروجی شود. انتخاب



نقطه هدف، در واقع ثابت شدن به صورت قائم هست، از شکل ۹ مشاهده می‌شود که تاریک‌ترین نقاط واقع در صفحه (زردترین نقاط) دارای بیشترین درجه تعلقی‌ها هستند و دلیل آن تلاش بیشتر برای هدایت پاندول حول این نقطه و قرار داشتن تعداد داده‌های آموزشی بیشتر حول این نقطه هدف هست. مسیرهای باریک استخراج شده برای زوایای انحراف کوچک، همگی خط راست هستند. داده‌های مناسب، داده‌های هستند که با اعمال نیرو به سیستم، زاویه و سرعت زاویه‌ای را به حول و حوش صفر می‌رسانند. نقطه هدف که در واقع ثابت شدن به صورت قائم است، تاریک‌ترین نقطه و یا نقطه با بیشترین درجه اعتقاد هست.

## ۵- شبیه‌سازی

در این بخش جهت ارزیابی الگوریتم ارائه شده شبیه‌سازی‌هایی به- ازای ورودی‌های مختلف صورت پذیرفته است. تمامی شبیه‌سازی‌ها در برنامه متلب 2017a و با یک کامپیوتر شخصی با پردازنده Core i5، 2.4 GHz و 4GB رم صورت گرفته شده است. در شبیه‌سازی‌ها تعداد سطوح کوانتیزاسیون محورها برابر با ۲۵۶ در نظر گرفته شده است. برای شبیه‌سازی باید از یک سیستم تحت کنترل و یا حتی سیستم بدون کنترلر دیتا جمع‌آوری کنیم ولی سیستم پاندول واقعی موجود نیست و به همین دلیل جمع‌آوری داده‌ها از شبیه‌سازی سیستم بدست آورده می‌شود. برای جمع‌آوری داده از معادله تفاضلی بهره گرفتیم و این رابطه به‌ازای مقادیر اولیه تصادفی اجرا شده است (در کنترلر غیر کلاسیک نیازی به دانستن مقدار F نداریم). در این رابطه،  $y$  بر حسب سانتی‌متر و  $\theta$  بر حسب رادیان هست.

$$\theta[n] - 2.2613\theta[n-1] + \theta[n-2] = 2.6633[n] - 5.3266[n-1] + 2.6633y[n-2] \quad (11)$$

برای رسیدن به رابطه فوق از مقادیر پارامترهای جدول 1 بهره گرفته شده است. ضمناً تغییرات  $\theta = [-0.23, 0.23]$  در این بازه قرار دارد و این تغییرات بر حسب رادیان است (یا اینکه تغییرات مثبت و منفی ۱۳ درجه می باشد).

جدول (۱): مشخصات پاندول استفاده شده و شتاب

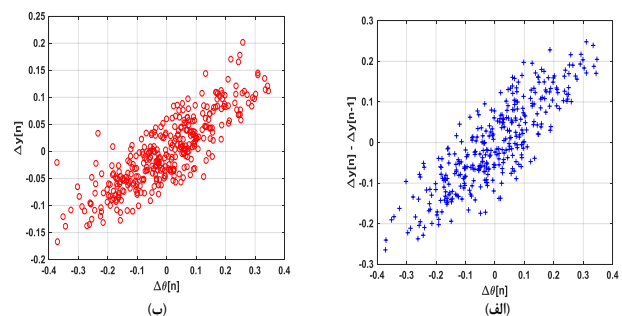
$$g = 9.81 \frac{m}{s^2} \text{ گرانشی}$$

| پارامترها | طول پاندول | جرم گاری    | جرم پاندول  | نمونه برداری |
|-----------|------------|-------------|-------------|--------------|
| مقادیر    | $l = 0.6m$ | $m = 0.2Kg$ | $m = 0.2Kg$ | $f_s = 10Hz$ |

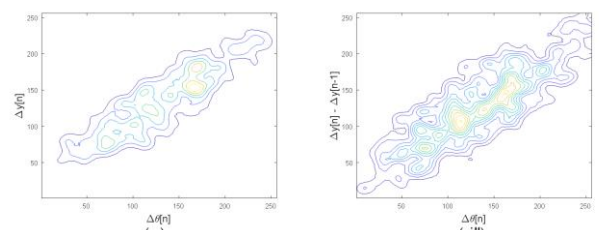
برای بررسی کارایی به مقایسه الگوریتم در حالت‌های بدون اثردهی داده‌های منفی و با اثردهی داده‌های منفی پرداخته شده است. ساختار کلی کنترلر در شکل ۱۰ آورده شده است. با توجه به تغییرات  $\theta$  در

داده‌های منفی به این صورت است که بعد شروع حرکت تصادفی سیستم مورد کنترل، حرکت‌هایی که نسبت به حرکت قبلی عملکرد بدتری دارند، یعنی سبب دور شدن پاندول از نقطه تعادل می‌شوند (باعث دور شدن سیستم از هدف اصلی) به‌عنوان داده منفی انتخاب می‌شوند و برعکس این انتخاب به‌عنوان داده‌های مثبت در نظر گرفته می‌شوند.

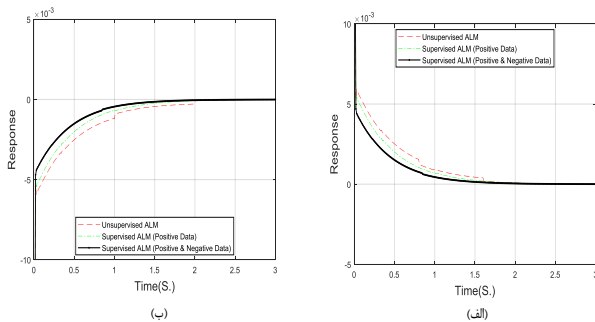
در روند طراحی کنترلر با استفاده از الگوریتم ALM، به این صورت عمل خواهیم کرد که با ترکیب زیرسیستم‌های تک- ورودی تک- خروجی، تلاش در تخمین وضعیت بعدی سیستم را داریم و پس از آن با اعمال ورودی برای تغییر وضعیت به سمت دلخواه، سیستم را به سمت نقطه کار نزدیک می‌کنیم. در روش ALM خروجی سیستم است که باید به پایداری برسد. در شکل ۸) کمترین پراکندگی در صفحات IDS در فرآیند مدل‌سازی انتخاب شده است و همچنین در شکل ۹) نقطه تعادل پررنگ ترین نقطه در تمامی صفحات IDS است. هنگامی که یک آپراتور انسانی به بهبود قدرت خود در کنترل یک پدیده می‌پردازد، برخی بازه‌های پارامترها یا تغییرات آن‌ها را به‌عنوان مقادیر بحرانی شناسایی کرده و برخی بازه‌های دیگر را غیرمهم تلقی می‌کند و سپس در مناطق بحرانی با دقت‌تر عمل می‌کند. همانند شکل ۸) متغیرهای اختلاف زاویه فعلی و قبلی (تغییرات زاویه)، تغییرات جابجایی ارا به و اختلاف تغییرات جابجایی فعلی و قبلی به‌عنوان مؤثرترین ورودی‌ها انتخاب شده است. در شکل ۹) صفحات IDS مرتبط با این متغیرها آورده شده است. البته در [۲۳] به بررسی و مدل‌سازی داده‌های مؤثر (انتخاب زیرسیستم‌هایی با کمترین پراکندگی) پرداخته شده است.



شکل ۸): نمایش شکل‌های الف) و ب) بیانگر، کمترین پراکندگی در صفحات IDS در فرآیند مدل‌سازی



شکل ۹): نمایش شکل‌های الف) و ب)، صفحات IDS به‌ازای پخش قطرات جوهر در صفحات با رزولوشن ۲۵۶ =  $R_{mix} = R_{my}$  و شعاع ۱۶



شکل (۱۲): (الف) پاسخ مرتبط با کنترلرها به ازای ورودی صفر، مقدار اولیه عبارت است از  $\theta = 0.21$ ، (ب) پاسخ مرتبط با کنترلرها به ازای ورودی صفر، مقدار اولیه عبارت است از  $\theta = -0.22$

برای بررسی اعتبار روش پیشنهادی به مقایسه سه کنترلر فوق در پاسخ به ورودی صفر و ورودی سینوسی پرداخته شده است. ورودی کنترلرها، تغییرات زاویه ایجاد شده برای پاندول هست و در نهایت خروجی کنترلرها تغییرات جابجایی برای دستیابی به این زاویه‌ها خواهد بود. در حالت کلی ۳۷۳ داده آموزشی استفاده شده است که ۲۸ درصد این داده‌ها (۱۰۴ داده) تحت عنوان داده‌های منفی هستند. نتایج به ازای ورودی صفر و ورودی سینوسی به ترتیب در شکل‌های (۱۲) و (۱۳) آورده شده است. برای بررسی دقت الگوریتم و خطای مدل سازی از معیار  $FVU$  (Fraction of Variance Unexplained) استفاده شده است که در مقاله [۱۴] از آن بهره گرفته و نحوه محاسبه آن با توجه به رابطه زیر است.

$$FVU = \frac{\sum_{i=1}^k (\hat{y}(x_i) - y(x_i))^2}{\sum_{i=1}^k (y(x_i) - \bar{y})^2} \quad (12)$$

$$\bar{y} = \frac{1}{k} \sum_{i=1}^k y(x_i) \quad (13)$$

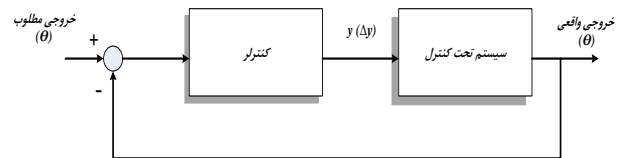
در رابطه فوق  $\hat{y}$  خروجی حاصل از مدل بوده و  $k$  بیانگر تعداد داده‌های تست است. این معیار رابطه مستقیم با  $MSE$  داشته، به طوری که با صفر شدن آن  $MSE$  نیز صفر می‌شود. شبیه‌سازی مرتبط در جدول ۲ زیر آورده شده است.

جدول (۲): مقایسه  $FVU$  بدست آمده از شبیه‌سازی به ازای ورودی سینوسی

| روش   | Unsupervised ALM    | Supervised ALM (Pos. data) | Supervised ALM (Pos. & Neg. data) |
|-------|---------------------|----------------------------|-----------------------------------|
| $FVU$ | $0.0178 \pm 0.0067$ | $0.0168 \pm 0.0054$        | $0.0143 \pm 0.0037$               |

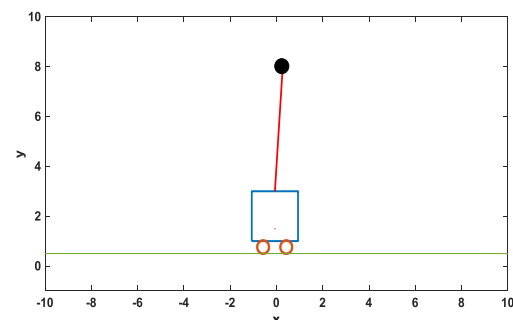
همان‌طور که مشاهده می‌شود، اگرچه تعداد نمونه‌های دسترس کنترلر ۲ از ۱ کمتر است، اما نتایج آن بهتر است. همچنین اگرچه نمونه‌های دسترس کنترلرهای ۱ و ۳ مساوی است، اما نتایج کنترلر ۳

خروجی Plant با تغییرات مناسب برای  $y$ ، اقدام به کنترل پاندول معکوس می‌نماید.



شکل (۱۰): نوع ورودی و خروجی در کنترلر طراحی شده و سیستم کنترلی

نمایش گرافیکی از پاندول معکوس شبیه‌سازی شده در محیط متلب در شکل (۱۱) آورده شده است.



شکل (۱۱): نمایش گرافیکی پاندول معکوس در محیط متلب

کنترلر طراحی شده قصد دارد که تغییرات  $\Delta\theta$  را به صفر برساند. برای شبیه‌سازی سه حالت مختلف در نظر گرفته شده است. ۱) کنترلر اول: با در نظر گرفتن همه داده‌ها تحت عنوان داده‌های مثبت بدست آمده که فرآیند آن مشابه فرآیند یادگیری بدون نظارت است.

۲) کنترلر دوم: داده‌هایی را تحت عنوان حرکت‌های مناسب برای فرایند کنترل در نظر می‌گیرد (با نظارت). تقسیم‌بندی داده‌ها از طریق جدا کردن آن‌ها تحت عنوان حرکت‌های مناسب و غیرمناسب صورت گرفته است. حرکت مناسب حرکتی است که منجر به انحرافی کمتر از مقدار اولیه در خروجی شود.

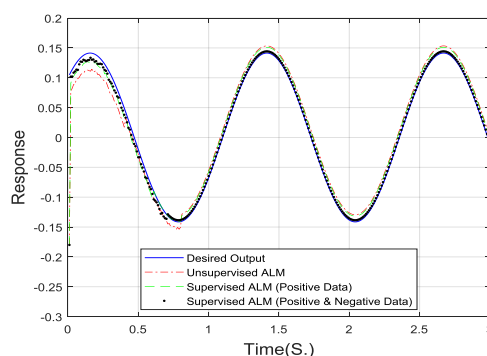
۳) کنترلر سوم: از طریق تقسیم‌بندی داده‌ها تحت عنوان داده‌های مثبت و منفی و در نظر گرفتن همه داده‌ها حاصل شده است (با نظارت). داده‌ای مثبت نمایانگر حرکت مناسب و داده‌ای منفی نمایانگر حرکت غیرمناسب است. نکته دیگری که در کنترلر سوم به آن توجه شده است، در نظر گرفتن اعتقاد به درستی یک تجربه است. اعتقاد به درجه درستی تجربه، یعنی اینکه به‌طور مثال در الگوریتم پخش قطره جوهر، آن قطره جوهر را در عدد یک یا عدد کوچک‌تر از یک ضرب کنیم. داده‌هایی که به نقطه تعادل نزدیک‌تر هستند، در ضرب بزرگ‌تری ضرب می‌شوند.

اجراء می‌شود. در این مقاله به طراحی سیستم کنترلی مبتنی بر ALM برای سیستم پاندول معکوس و تعادل آن پرداخته شده است. با بکارگیری هم زمان داده‌های مثبت و داده‌های منفی فضای تجربه مدل-سازی شد و با توجه به نتایج بدست آمده ملاحظه شده است که کنترلهای فازی طراحی شده، سیستم را با سرعت بیشتری به تعادل می-رساند. بنابراین داده‌های منفی توانسته‌اند که نقش مؤثری در یادگیری ایفا نمایند. به‌عنوان یکی از کارهای آینده، کنترل پاندول معکوس دو لینکه مورد بررسی قرار خواهد گرفت.

## مراجع

- [1] Boubaker, O. The inverted pendulum: A fundamental benchmark in control theory and robotics. in International conference on education and e-learning innovations. 2012. IEEE.
- [2] Bugeja, M. Non-linear swing-up and stabilizing control of an inverted pendulum system. in The IEEE Region 8 EUROCON 2003. Computer as a Tool. 2003. IEEE.
- [3] Prasad, L.B., B. Tyagi, and H.O. Gupta. Modelling and simulation for optimal control of nonlinear inverted pendulum dynamical system using PID controller and LQR. in 2012 Sixth Asia Modelling Symposium. 2012. IEEE.
- [4] Waszak, M. and R. Langowski, An automatic self-tuning control system design for an inverted pendulum. IEEE Access, 2020. 8: p. 26726-26738.
- [5] Jain, A., et al., Real-Time Swing-up Control of Non-Linear Inverted Pendulum using Lyapunov based Optimized Fuzzy Logic Control. IEEE Access, 2021. 9: p. 50715-50726.
- [6] Chaturvedi, D., T. Qamar, and M. Gupta, Neuro-Control of Inverted Pendulum, in Soft Computing for Problem Solving, 2021, Springer. p. 73-89.
- [7] Rithirun, C., A. Charean, and W. Sawaengsinsakikit. Comparison Between PID Control and Fuzzy PID Control on Invert Pendulum System. in 2021 9th International Electrical Engineering Congress (iEECON). 2021. IEEE.
- [8] Veysi, m., J. Aghaei, and m.r. soltanpour, Disturbance Rejection Stable Fuzzy Controller Design for a Class of Nonlinear Electromechanical systems, in presence of Parametric Uncertainties. Journal of Iranian Association of Electrical and Electronics Engineers, 2019. 16(3): p. 111-124.
- [9] Fallah Ghavidel, H., S.M. Mousavi Gzafroudi, and R. Asad, Thrust Control of BLDC Thruster Motors by Observer-Based Robust Adaptive Fuzzy Control. Journal of Iranian Association of Electrical and Electronics Engineers, 2020. 17(3): p. 109-118.
- [10] Shouraki, S.B. and N. Honda. Fuzzy controller design by an active learning method. in 31th symposium of intelligent control. Tokyo, Japan. 1998.
- [11] 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.

به طور قابل توجهی از دو کنترلر دیگر بهتر است. بنابراین وارد نمودن داده‌های منفی سبب عملکرد بهتر الگوریتم شده است و تعداد حرکت-های مورد نیاز برای ثابت کردن پاندول پس از تقسیم‌بندی داده‌ها تحت عنوان حرکت‌های مناسب و غیرمناسب، کاهش داشته است. تعداد شکست‌ها در هر کدام از صفحات IDS برابر با ۳ بوده و در نتیجه تعداد قوانین استفاده شده برابر با ۹ است. ضمن اینکه به‌کارگیری داده مثبت و هم داده منفی در پیاده‌سازی کنترلر ALM دارای سرعت بیشتر نسبت به دو مورد قبلی است. استفاده از داده منفی در فرآیند یادگیری، باعث افزایش سرعت یادگیری و فرآیند کنترل خواهد شد و فرآیند یادگیری با تعداد نمونه کمتری قابل انجام است. در انتها می-توان بیان داشت که، مثبت یا منفی بودن داده‌های آموزشی، به علت اینکه از محیط فیدبک می‌گیریم و این فیدبک در نحوه پخش داده‌ها تأثیرگذار است، سبب ایجاد وفقی بودن می‌شود. عملاً بهره‌گیری از داده مثبت و منفی مشابهت بسیاری با هرس کردن داده‌ها داشته و همانند ذهن انسان که از تجارب مثبت و منفی در تصمیم‌گیری‌های آینده بهره می‌گیرد است.



شکل (۱۳): پاسخ مرتبط با کنترلرها به‌ازای ورودی سینوسی، مقدار

$$\theta = -0.18 \text{ اولیه عبارت است از}$$

## ۶- نتیجه‌گیری

تجربه کار روی مسئله پاندول معکوس در دو دسته کلاسیک و فازی تقسیم‌بندی می‌شود. در کنترل پاندول معکوس، باید برای متعادل نگه داشتن آن، از حضور در برخی نواحی ممنوعه دوری شود. کنترلر طراحی شده مبتنی بر آپراتور IDS این امکان را فراهم می‌آورد که امکان معکوس‌پذیری مدل سیستم در مدل‌سازی را داشته باشیم. در این مقاله بررسی اثرات داده‌های منفی در فرآیند یادگیری بهره گرفته شده است. داده منفی، به داده‌های نسبت داده می‌شود که تأثیر منفی دارند و سیستم باید از این داده‌ها دوری کند. داده‌های منفی سبب می‌شود که بخشی نامتقارن در فضای IDS ایجاد شده، یعنی در جهتی که داده مناسب مشاهده شده، شعاع پخش‌شدگی بیشتر و در جهاتی که داده مناسب مشاهده نشده، شعاع پخش‌شدگی کمتر باشد. با این اقدام، سیاستی مشابه سیاست پاداش و جزاء در روش‌های یادگیری تقویتی



- [12] Sakurai, Y., A study of the learning control method using PBALM-a nonlinear modeling method. PhD, The University of Electro-Communications, Tokyo, 2005.
- [13] Sagha, H., I.E.P. Afrakoti, and S. Bagherishouraki, Actor-critic-based ink drop spread as an intelligent controller. *Turkish Journal of Electrical Engineering and Computer Science*, 2013. **21**(4): p. 1015-1034.
- [14] Murakami, M. and N. Honda, A study on the modeling ability of the IDS method: A soft computing technique using pattern-based information processing. *International journal of approximate reasoning*, 2007. **45**(3): p. 470-487.
- [15] 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.
- [16] 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.
- [17] 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.
- [18] 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.
- [19] 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.
- [۲۰] سجاد حق زاد کلیدبری، سعید باقری شورکی، ایمان اسمعیلی  
پابین افراکتی، ارائه اپراتور جدید جایگزین پخش قطره جوهر در روش  
یادگیری فعال. *مجله مهندسی برق دانشگاه تبریز*, ۲۰۱۹. **۴۹**(۳): p. 1055-1066.
- [21] 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.
- [22] Ghasemi, V., M. Javadian, and S. Bagheri Shouraki, High-Dimensional Unsupervised Active Learning Method. *Journal of AI and Data Mining*, 2020. **8**(3): p. 39. ۴۰۷-۱.
- [23] Shouraki, S.B., A novel fuzzy approach to modeling and control and its hardware implementation based on brain functionality and specifications. 2000.

---

زیر نویس‌ها:

<sup>1</sup> Active Learning Method (ALM)

<sup>2</sup> Ink Drop Spread (IDS)

<sup>3</sup> Reinforcement Learning (RL)

<sup>4</sup> Inverted Pendulum System (IPS)

<sup>5</sup> Narrow Path (NP)

<sup>6</sup> Spread (SP)