

دسته بندی اهداف سوناری با استفاده از شبکه های عصبی مصنوعی آموزش دیده مبتنی بر جغرافیای زیستی

سید محمدرضا موسوی^{۱*}، محمد خویشه^۲، فلاح محمدزاده^۳، هومان علائیان^۴

۱- استاد، دانشگاه علم و صنعت ایران، ۲- دانشجوی دکتری، دانشگاه علم و صنعت ایران، ۳- مربی، دانشگاه علوم دریایی امام خمینی (ره) نوشهر

۴- دانشجوی کارشناسی مهندسی برق، دانشگاه علم و صنعت ایران،

(دریافت: ۹۵/۰۷/۰۵، پذیرش: ۹۶/۰۳/۰۶)

چکیده

هدف این مقاله، استفاده از الگوریتم بهینه شده مبتنی بر جغرافیای زیستی برای آموزش شبکه های عصبی چندلایه به منظور دسته بندی اهداف سوناری می باشد. به منظور آزمایش عملکرد این الگوریتم، دسته بندی کننده طراحی شده بوسیله دادگان سونار آزموده شده و نتایج بدست آمده با پنج الگوریتم تکاملی معروف مقایسه می شود. معیارهای مورد سنجش عبارتند از: سرعت همگرایی، احتمال گیر افتادن در بهینه های محلی و دقت دسته بندی. نتایج نشان می دهند که آموزش شبکه های عصبی چندلایه با استفاده از الگوریتم بهینه شده مبتنی بر جغرافیای زیستی نسبت به الگوریتم های تکاملی موجود و الگوریتم پس انتشار، در تمام زمینه ها نتایج بهتر و یا قابل مقایسه ای را ارائه می کند.

واژه های کلیدی: شبکه های عصبی، سونار، الگوریتم بهینه شده مبتنی بر جغرافیای زیستی، الگوریتم تکاملی

۱- مقدمه

وابسته به آزمایشات گسترده بوده و در نتیجه هزینه، زمان و تجهیزات بسیار زیادی را احتیاج دارند. از طرف دیگر، با توجه به این که مشخصات انتشار صوت در هر نقطه از دریا و در هر باز زمانی از شب و روز متفاوت است، دسته بندی کننده ای باید انتخاب شود که هم به آزمایشات کم تر وابسته باشد و هم قابلیت تطبیق پذیری بالایی داشته باشد. از این رو، در این مقاله شبکه های عصبی مصنوعی برای دسته بندی هدف واقعی و اکوی بدون هدف انتخاب می شوند. این گونه شبکه ها می توانند با یک مجموع دادگان محدود آموزش ببینند و در سناریوهای آزمایش مختلف به کار برده شوند [۳]. صرف نظر از کاربردها، توانایی متمایز شبکه های عصبی مصنوعی چندلایه، یادگیری می باشد [۴]. یادگیری به این معنی است که این شبکه ها همانند مغز انسان می توانند از یک تجربه یا آزمایش یاد بگیرند. این ویژگی (یادگیری) بخش ضروری همه شبکه های عصبی است که ممکن است به دو نوع تقسیم گردد: یادگیری با نظارت^۳ [۵] و یادگیری بدون نظارت^۴ [۶]. برای شبکه های عصبی مصنوعی چندلایه (در بیش تر کاربردها)، از الگوریتم های پس انتشار^۵ بهینه شده [۷] و یا استاندارد [۸]، به عنوان روش یادگیری استفاده می گردد که از خانواده یادگیری با نظارت می باشند. الگوریتم پس انتشار بر مبنای گرادینان است که اشکالاتی هم چون هم گرایی آهسته [۹] و

اهداف آشکار سازی شده توسط سونار عبارتند از: هدف واقعی، نویز، طنین و اکوی بدون هدف. نویز دارای انواع مختلفی مانند نویز حرارتی، محیطی، قاره ای، لنگرگاه و غیره می باشد که چون از جنس صوت ارسالی توسط سونار نمی باشد، تمایز آن از هدف واقعی آسان است. اکوهای حاصل از برخورد پینگ ارسالی سونار با کف و سطح دریا را طنین گویند [۱]. با توجه به این که اکوهای حاصل از طنین دارای یک دامنه یکسان و همگن می باشند، علی رغم این که از جنس پینگ ارسالی هستند، تمایز آن ها از هدف واقعی آسان است. هنگامی که جنس بستر دریا دارای تغییرات زیادی است و در ابعاد کوچک، بستر با جنس های مختلفی وجود دارد، اکوهای برگشتی از بستر دارای ویژگی های هدف گونه خواهند بود، به صورتی که حتی تابع چگالی احتمال هدف واقعی و بستر بسیار باهم شبیه خواهند بود. این گونه اهداف کاذب را اکوی بدون هدف گویند [۲]. طبقه بندی اکوی بدون هدف و اهداف واقعی به خاطر تشابه بسیار زیاد اکوهای برگشتی از آن ها، کار بسیار دشواری است. دسته بندی کننده های متعارف آماری (که بیش تر بر مبنای نظریه بیزن^۱ استوار هستند) به دلیل نیازمندی به محاسبه دانش قبلی و تابع چگالی احتمال^۲ (PDF)،

3- Supervised Learning

4- Unsupervised Learning

5- Back-Propagation (BP) Algorithm

* رایانامه نویسنده مسئول: m_mosavi@iust.ac.ir

1- Bayesian Theorem

2 - Probability Density Function

توجه به ماهیت تصادفی روش‌های فراابتکاری، مسئله‌ای چالش برانگیز می‌باشد. یکی از الگوریتم‌های بسیار توانا در این زمینه بهینه‌ساز مبتنی بر جغرافیایی زیستی^۸ (BBO) می‌باشد که با تعداد پارامترهای کم توانایی تطبیق برای مجموعه دادگان‌های مختلف را دارا می‌باشد. به عبارت دیگر با تنظیم خیلی ساده پارامترهای این الگوریتم، می‌توان الگوریتم بهینه‌ساز جدیدی را تولید کرد که برای آن مسأله خاص بهترین جواب ممکن را تولید کند. در این مقاله برای اولین بار الگوریتم BBO را برای دادگان سونار تغییر داده و نرخ‌های جدیدی برای مهاجرت به داخل^۹ و خارج^{۱۰} معرفی می‌کنیم و هم‌چنین الگوریتم جدید را با نام IBBO^{۱۱} برای آموزش شبکه عصبی MLP به کار برده و توسط آن دسته‌بندی‌کننده جدیدی با نام MLP-IBBO طراحی می‌نمایم.

در این مقاله، شبکه عصبی MLP توسط الگوریتم IBBO به منظور دسته‌بندی دادگان سونار (شامل هدف واقعی و هدف کاذب) آموزش داده شده است. دلیل استفاده از شبکه MLP و الگوریتم آموزشی IBBO به شرح ذیل می‌باشد:

- شبکه MLP به دلیل استفاده از توابع سیگموئید، توانایی منحصر به فردی در کار با دادگانی که به‌طور خطی قابلیت تفکیک ندارند را دارا می‌باشد. از طرفی دادگان سونار به‌صورت خطی قابل تفکیک نیستند و نیازمند دسته‌بندی‌کننده‌ای با ابعاد بالا می‌باشند.
- اکوی بدون هدف و هدف واقعی دارای ویژگی‌های بسیار شبیه هم می‌باشند. در نتیجه باید الگوریتمی انتخاب شود که فضای جستجو را به‌طور کامل اکتشاف کند.
- نقطه قوت الگوریتم IBBO نسبت به الگوریتم‌های فراابتکاری دیگر، قدرت اکتشاف فوق‌العاده آن می‌باشد. در مقایسه با روش‌های هوش گروهی^{۱۲} (از قبیل روش بهینه‌سازی ازدحام ذرات و روش بهینه‌سازی کلونی مورچه‌ها)، عملگر جهش الگوریتم ژنتیک توانایی اکتشاف بیش‌تری را فراهم می‌کند. این پتانسیل موجب می‌شود تا بهینه‌سازی مبتنی بر جغرافیایی زیستی عملکرد بهتری در آموزش شبکه‌های عصبی مصنوعی چندلایه نسبت به تکنیک‌های هوش گروهی داشته باشد. به‌علاوه، در روش بهینه‌سازی مبتنی بر جغرافیایی زیستی، داشتن ثابت‌های جهش مختلف برای هر فرد در یک جمعیت، نسبت به الگوریتم ژنتیک که فقط یک عملگر جهش برای کل جمعیت دارد، به ایجاد عملکرد بهتر کمک می‌کند.
- نقطه ضعف الگوریتم BBO سرعت کم در فاز بهره‌برداری است. از طرف دیگر در بسیاری از کاربردهای سونار فعال (از

به‌کارگیری در یک محدوده کوچک [۱۰] را دارد و بنابراین برای کاربردهای عملی قابل اعتماد نیست.

هدف نهایی فرآیند یادگیری در شبکه‌های عصبی، پیدا کردن بهترین ترکیب از یال‌های وزن‌دار و حد آستانه^۱ آنها است به‌طوری که در آموزش شبکه و نمونه‌های آزمون، کم‌ترین مقدار خطا را داشته باشیم. اغلب خطای شبکه عصبی چندلایه، برای مدت زیادی در زمان فرآیند یادگیری، بزرگ خواهد بود و الگوریتم یادگیری آن را به سمت کم‌شدن هدایت می‌کند. این مسأله در فرآیندهای یادگیری مبتنی بر گرادیان، مثل الگوریتم پس‌انتشار، مشترک است. هم‌چنین هم‌گرایی الگوریتم پس‌انتشار به مقدار خیلی زیاد به مقادیر اولیه نرخ یادگیری و اندازه حرکت وابسته است. مقادیر نامناسب این متغیرها می‌تواند حتی سبب واگرایی الگوریتم گردد. مطالعات بسیار زیادی برای حل این مشکل الگوریتم پس‌انتشار انجام گرفته است [۱۱]. اما نتایج مورد انتظار به دست نیامده و هر روش فقط تأثیرات جانبی خودش را داشته است. مقاله [۱۲] نشان می‌دهد که الگوریتم‌های جستجوی ابتکاری یا اکتشافی^۲ می‌تواند جایگزین الگوریتم‌های یادگیری مبتنی بر گرادیان باشد، زیرا ماهیت تصادفی این الگوریتم‌ها اجازه می‌دهد تا حداقل خطای بهتری نسبت به روش‌های مبتنی بر گرادیان داشته باشیم ولی اکتشاف تمام حالات، پیچیدگی زمانی و مکانی شبکه را افزایش می‌دهد که برای مسائل با ابعاد بالا این مشکل شدیدتر می‌شود [۱۳].

علاوه بر روش‌های مبتنی بر مشتق و روش‌های اکتشافی، در سال‌های اخیر روش‌های فراابتکاری گوناگونی از قبیل بهینه‌سازی گروهی ذرات^۳ [۱۴]، الگوریتم ژنتیک^۴ [۱۵]، الگوریتم اجتماع مورچه‌ها^۵ [۱۶] و الگوریتم‌های تکامل‌پذیر^۶ [۱۷] برای آموزش شبکه‌های عصبی به کار گرفته شده است. با استفاده از قضیه "هیچ نهاری مجانی نیست"^۷، ثابت شده است که الگوریتم‌های جستجوی اکتشافی الزاماً بهترین جواب را برای مسائل بهینه‌سازی تولید نمی‌کنند [۱۱ و ۱۹-۱۸]. این قضیه از یک طرف و مشکلات روش‌های مبتنی بر گرادیان از طرف دیگر، بسیاری از محققان را بر آن داشته است که درخصوص تأثیر الگوریتم‌های فراابتکاری متفاوت در یادگیری شبکه‌های عصبی مصنوعی تلاش‌های بسیاری را انجام دهند [۲۷-۲۰]. صرف‌نظر از تفاوت‌های بین روش‌های فراابتکاری مختلف، یک ویژگی مشترک بین آنها تقسیم فرآیند جستجو به دو مرحله شناسایی و بهره‌برداری است [۲۸-۲۹]. پیدا کردن یک موازنه مناسب بین این دو مرحله با

- 1- Bias
- 2- Heuristic Optimization Methods
- 3- Particle Swarm Optimization (PSO)
- 4- Genetic Algorithm (GA)
- 5- Ant Colony Optimization (ACO)
- 6- Evolutionary Strategies (ES)
- 7- No Free Lunch Theorem (NFL)

- 8- Biogeography based Optimization Algorithm (BBO)
- 9- Immigration
- 10- Emigration
- 11- Improved-BBO
- 12- Swarm Intelligence (SI)

$$S_j = \text{sigmoid}(s_j) = \frac{1}{(1 + \exp(-s_j))}, \quad j=1,2,\dots,h \quad (2)$$

می‌توان بعد از محاسبه مقدار گره‌های پنهان، خروجی‌های نهایی را به صورت زیر تعریف نمود:

$$o_k = \sum_{j=1}^h (W_{jk} \cdot S_j) - \theta'_k, \quad k=1,2,\dots,m \quad (3)$$

$$O_k = \text{sigmoid}(o_k) = \frac{1}{(1 + \exp(-o_k))}, \quad k=1,2,\dots,m \quad (4)$$

که در آن، W_{jk} مبین وزن یال متصل‌کننده گره j -ام (در لایه پنهان) به گره k -ام (در لایه خروجی) و θ'_k مبین بایاس گره k -ام (در لایه خروجی) است. مهم‌ترین بخش‌های شبکه‌های عصبی چندلایه، وزن یال‌ها و بایاس گره‌ها می‌باشد. همان‌طور که در روابط بالا مشاهده گردید، وزن یال‌ها و بایاس‌ها، مقدار خروجی نهایی را تعریف کردند. آموزش یک شبکه عصبی چندلایه، شامل پیدا کردن بهترین مقدار برای وزن یال‌ها و بایاس‌ها، به منظور رسیدن به مقدار مطلوب خروجی در ازای ورودی‌های مشخص است.

۳- الگوریتم بهینه‌سازی مبتنی بر جغرافیای زیستی

الگوریتم بهینه‌سازی مبتنی بر جغرافیای زیستی برای اولین بار در سال ۲۰۰۸ توسط سیمون^۱ [۳۰] پیشنهاد گردید. ایده اصلی این الگوریتم از رشته‌ای در زیست‌شناسی که درباره طرز انتشار و پخش حیوانات و نباتات (در زمان و مکان) بحث می‌کند، الهام گرفته شده است. در این الگوریتم، اکوسیستم‌های متفاوت (محل سکونت یا قلمروها)، برای یافتن ارتباط بین گونه‌های مختلف (ساکنین) بر حسب مهاجرت به بیرون، مهاجرت به داخل و جهش، بررسی می‌شوند. تکامل اکوسیستم‌ها با در نظر گرفتن انواع گونه‌های مختلف و تأثیر مهاجرت و جهش برای رسیدن به یک شرایط پایدار، زیربنای الگوریتم بهینه‌سازی مبتنی بر جغرافیای زیستی می‌باشد.

همانند الگوریتم ژنتیک، الگوریتم بهینه‌سازی مبتنی بر جغرافیای زیستی، از تعدادی عوامل جستجو به نام محل‌های سکونت^۲ استفاده می‌کند. این محل‌های سکونت، مشابه کروموزوم‌ها در الگوریتم ژنتیک هستند. الگوریتم بهینه‌سازی مبتنی بر جغرافیای زیستی، هر محل سکونت را به صورت

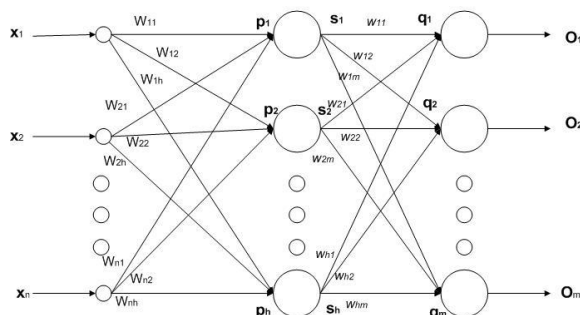
قبیل کاربردهای نظامی) نیاز به پردازش‌های بلادرنگ است و پیچیدگی زمانی یک نقطه ضعف بزرگ به‌شمار می‌آید. این ضعف توسط نرخ‌های جدید مهاجرت جبران می‌شود. سازمان‌دهی مقاله بدین صورت می‌باشد که بخش ۲ به معرفی شبکه‌های عصبی مصنوعی چندلایه خواهد پرداخت. بخش ۳ به بحث در خصوص کلیات روش BBO، تنظیم نرخ‌های مهاجرت و نحوه تولید IBBO پرداخته است. شیوه اعمال IBBO به عنوان یک الگوریتم آموزش تکاملی در شبکه‌های عصبی مصنوعی چندلایه، در بخش ۴ توصیف گردیده است. نتایج در بخش ۵ مورد بحث قرار خواهند گرفت. در نهایت در بخش ۶ نتیجه‌گیری و محورهای مطالعاتی که می‌توان در خصوص آن‌ها به مطالعه پرداخت، بیان خواهد شد.

۲- شبکه‌های عصبی چندلایه

شکل (۱) یک شبکه عصبی چندلایه (۳ لایه) را نشان می‌دهد که در آن، n مبین تعداد گره‌های ورودی، h مبین تعداد گره‌های پنهان و m مبین تعداد گره‌های خروجی است. همان‌طور که ملاحظه می‌شود، اتصال‌های یک طرفه بین گره‌های شبکه عصبی چندلایه که از خانواده شبکه‌های عصبی MLP است، وجود دارد. خروجی شبکه عصبی چندلایه به صورت رابطه (۱) محاسبه می‌گردد:

$$s_j = \sum_{i=1}^n (W_{ij} \cdot X_i) - \theta_j, \quad j=1,2,\dots,h \quad (1)$$

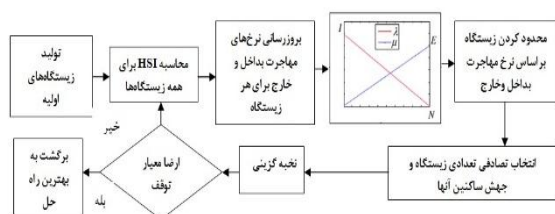
که در آن، n مبین تعداد گره‌های ورودی، W_{ij} مبین وزن یال متصل‌کننده گره i -ام (در لایه ورودی) به گره j -ام (در لایه پنهان)، θ_j مبین بایاس گره j -ام (در لایه پنهان) و X_i مبین ورودی به گره i -ام (در لایه ورودی) است. خروجی هر گره پنهان با استفاده از یک تابع سیگموئید و به صورت رابطه (۲) به دست می‌آید:



شکل (۱): یک شبکه عصبی مصنوعی چندلایه با یک لایه پنهان

1- Simon
2- Habitat

داده شده است. با استفاده از این شکل می توان فهمید که تعداد زیادی از ساکنین، در یک زمان با احتمال زیاد به بیرون مهاجرت می کنند و با احتمال کم، به داخل مهاجرت می نمایند [۳۰].



شکل (۲): مراحل عمومی الگوریتم بهینه سازی مبتنی بر جغرافیای زیستی

سومین مولفه الگوریتم بهینه سازی مبتنی بر جغرافیای زیستی، جهش است که قدرت اکتشاف الگوریتم بهینه سازی مبتنی بر جغرافیای زیستی را بهبود داده و تنوع محل های سکونت را حفظ می کند. این مولفه به صورت رابطه (۷) تعریف می شود:

$$m_n = M \times \left(1 - \frac{p_n}{p_{\max}}\right) \quad (7)$$

که در آن، M مبین مقدار اولیه جهش است که به وسیله کاربر تعریف می گردد. p_n مبین احتمال جهش n -امین محل های سکونت و p_{\max} مبین بیش ترین مقدار p_n است و به صورت رابطه (۸) تعریف می گردد:

$$p_{\max} = \arg \max(p_n) \quad (8)$$

در بلوک دیاگرام شکل (۲) مراحل کلی الگوریتم بهینه سازی مبتنی بر جغرافیای زیستی نشان داده شده است. این شکل نشان می دهد که الگوریتم بهینه سازی مبتنی بر جغرافیای زیستی با مجموعه ای از محل های سکونت تصادفی آغاز می گردد. بعد از محاسبه شاخص مناسب بودن زیستگاه^۳ برای هر محل سکونت، نرخ های مهاجرت به بیرون، مهاجرت به داخل و جهش نیز به روز (محاسبه) می شود. مطابق با این نرخ ها، ساکنین معمولی (غیرنخبه) مهاجرت کرده و یا جهش پیدا می کنند. تعدادی از بهترین محل های سکونت که از قبل تعریف شده اند برای تولید نسل های بعد در نظر گرفته می شوند. سرانجام الگوریتم بهینه سازی مبتنی بر جغرافیای زیستی با برآورده شدن شرایط پایانی، به پایان می رسد. ذکر این نکته لازم است که نخبه گرایی مانع از خراب شدن بهترین محل سکونت با مهاجرت به داخل می گردد. با انجام این کار، ما تعدادی از بهترین محل های سکونت را در هر تکرار حفظ می کنیم. در مورد مجموعه دادگان سنوار نرخ های خطی مهاجرت

بردارهایی از ساکنین^۱ (مشابه ژن ها در الگوریتم ژنتیک) در نظر می گیرد که متغیرهای مسأله را نشان می دهند. به علاوه، برای هر محل سکونت، شاخص مناسب بودن محل سکونت^۲ نیز تعریف می گردد. بالا بودن این شاخص به منزله داشتن شرایط بهتر می باشد. در هر زمان محل های سکونت بر اساس سه قانون اصلی به شرح زیر تعیین می گردند:

الف) ساکنینی که در مکان های با شاخص پایین اقامت دارند، بیش تر به مهاجرت به مکان هایی با شاخص بالاتر، تمایل دارند.

ب) ساکنینی که در مکان های با شاخص بالا اقامت دارند، تمایل بیش تری به جذب مهاجران از مکان های با شاخص پایین دارند.

ج) مکان ها بدون توجه به مقدار شاخص شان، باید به صورت تصادفی ساکنان شان را تغییر دهند.

در طبیعت، این پدیده موجب ایجاد تعادل میان اکوسیستم های مختلف می گردد. به عبارت دیگر، طبیعت به بهبود پایداری نواحی زیستی مختلف تمایل دارد. الگوریتم بهینه سازی مبتنی بر جغرافیای زیستی، این مفاهیم را برای بهبود شاخص همه مکان های زیستی به کار می برد که نتایج آن در استخراج راه حل تصادفی اولیه برای یک مسأله خاص مورد استفاده قرار می گیرد.

با انتخاب تصادفی مجموعه ای از محل های سکونت، الگوریتم بهینه سازی مبتنی بر جغرافیای زیستی آغاز می گردد. هر محل دارای n ساکن مختلف است که بر اساس متغیرهای یک مسأله خاص تعیین می گردند. به علاوه هر محل، نرخ های مهاجرت به بیرون، مهاجرت به داخل و جهش خاص خودش را دارد که از مکان های متمایز از نظر زیستی در طبیعت الگوبرداری شده است. مهاجرت به بیرون (μ_k) و مهاجرت به داخل (λ_k)، به عنوان توابعی از تعداد ساکنان آنان به صورت زیر تعریف می گردند:

$$\mu_k = \frac{E \times n}{N} \quad (5)$$

$$\lambda_k = I \times \frac{1-n}{N} \quad (6)$$

که در آن، n مبین تعداد ساکنین کنونی، N مبین حداکثر تعداد ساکنین مجاز که با استفاده از شاخص مناسب بودن محل سکونت افزایش می یابد (محل سکونت مناسب تر، تعداد ساکنین بیش تر)، E مبین بیش ترین نرخ مهاجرت به بیرون و I مبین بیش ترین نرخ مهاجرت به داخل است. در شکل (۲) بیش ترین نرخ مهاجرت به بیرون و بیش ترین نرخ مهاجرت به داخل، نشان

هم‌چنین $p \in \{2,3\}$ انتخاب شده است. در واقع ما توسط این روابط نرخ‌های مهاجرت موجودات را براساس میزان مقدار شاخص مناسب‌بودن محل سکونت (HSI) آن‌ها تنظیم می‌کنیم. همان‌گونه که در رابطه (۹) ملاحظه می‌شود، برای محل‌های سکونت با HSI پایین ($k \leq N/p$) نرخ مهاجرت به خارج به صورت لگاریتمی افزایشی است، درحالی‌که نرخ مهاجرت به داخل به صورت نمایی کاهش می‌یابد. عکس همین روابط را برای محل‌های سکونت با HSI بالا داریم. نحوه تنظیم این پارامترها هم براساس درک بالا از نحوه مهاجرت پرندگان و تاثیر مدل‌های ریاضی مختلف بر آن‌ها بوده است. به عبارت ساده‌تر، برای هر زیستگاه براساس مقدار HSI، نرخ‌های مهاجرت مختلف را در نظر گرفته و از سوی دیگر مدل ریاضی که این رفتار را به بهترین وجه توصیف کند با چند بار آزمایش به دست آورده‌ایم. به‌دست‌آوردن بهینه‌ترین مدل یکی از کارهایی است که می‌توان آن‌را به عنوان یکی از کارهای پیش‌رو در نظر گرفت.

در بخش بعدی، نخست الگوریتم IBBO به یک شبکه عصبی چندلایه اعمال و سپس با الگوریتم‌های BBO استاندارد، بهینه‌سازی ازدحام ذرات، ژنتیک، بهینه‌سازی کلونی مورچگان و جستجوی گرانشی روی دادگان سونار با ابعاد کاهش‌یافته، مقایسه می‌گردد.

۴- آموزش یک شبکه عصبی چندلایه با استفاده از الگوریتم بهینه‌سازی مبتنی بر جغرافیای زیستی

به‌طورکلی، از الگوریتم‌های تکاملی برای آموزش شبکه‌های عصبی چندلایه در سه حالت استفاده می‌شود. اولین روش عبارت است از به‌کارگیری شبکه‌های تکاملی به‌منظور یافتن ترکیب وزن یال‌ها و بایاس گره‌ها برای داشتن کم‌ترین مقدار خطا در یک شبکه عصبی چندلایه. دومین شیوه عبارت است از به‌کارگیری شبکه‌های تکاملی به‌منظور یافتن ساختار مناسب شبکه عصبی چندلایه در یک مسئله خاص و آخرین شیوه شامل به‌کارگیری شبکه‌های تکاملی به‌منظور یافتن پارامترهای الگوریتم یادگیری مبتنی بر گرادینان از قبیل نرخ یادگیری و اندازه حرکت است. در این مقاله، الگوریتم بهینه‌سازی مبتنی بر جغرافیای زیستی، با استفاده از شیوه نخست به یک شبکه عصبی چندلایه اعمال می‌گردد. به منظور طراحی یک الگوریتم آموزش‌دهنده برای شبکه‌های عصبی چندلایه، لازم است مراحل زیر انجام گردد:

نمی‌تواند نتایج خوبی را ارائه کند. بنابراین، در قسمت بعد با توجه به ویژگی‌های این مجموعه دادگان یک مدل غیرخطی ارائه می‌شود. نتایج نشان خواهد داد که این مدل دارای توانایی منحصربه‌فردی در افزایش دقت دسته‌بندی دادگان سونار با حفظ پیچیدگی شبکه است.

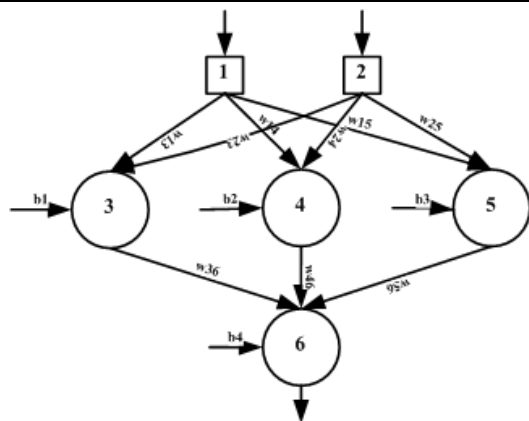
۳-۱ الگوریتم BBO با نرخ مهاجرت بهبودیافته (IBBO)

در BBO استاندارد، پدیده‌های مهاجرت به داخل و خارج به صورت خطی و طبق روابط (۵-۶) مدل شده‌اند. این در حالی است که با توجه به سازوکارهای مهاجرت در بین حیوانات و پیچیدگی‌های فراوان این پدیده، نمی‌توان پدیده مهاجرت را به صورت خطی مدل کرد. در مراجع گوناگون [۳۱-۳۲] برای کاربردهای مختلف از مدل‌های غیرخطی مختلف استفاده شده است. در بعضی کاربردها که نیاز به پردازش بلادرنگ می‌باشد [۳۳]، از مدل‌هایی که به سرعت هم‌گرا شوند، استفاده شده است و در کاربردهایی که رسیدن به جواب بهینه و قطعی اولویت می‌باشد [۳۴]، از مدل‌های توسعه‌ای استفاده شده است. مسئله دسته‌بندی دادگان سونار تا حدودی با مسائل مطرح شده قبلی متفاوت می‌باشد. بدین صورت که در این مجموعه دادگان خاص، از یک طرف، با توجه به شباهت بسیار زیاد هدف واقعی و هدف کاذب، به تابعی نیاز است که فضای جستجو را به خوبی اکتشاف کرده و تمام بهینه‌های محلی را شناسایی کند تا در بهینه‌های محلی گیر نکند. از سوی دیگر، با توجه به نیاز به پردازش بلادرنگ، به تابعی نیاز است که در فاز بهره‌برداری رفتار کاملاً متفاوتی داشته باشد و به سرعت هم‌گرا شود. با توجه به مباحث مطرح‌شده در این مقاله، مهاجرت به داخل و خارج توسط روابط (۹-۱۰) تعریف می‌شوند. بعد از اصلاح BBO متناسب با نوع مجموعه دادگان سونار، از این پس در این مقاله این الگوریتم را IBBO^۱ نام‌گذاری می‌کنیم. در ادامه مقاله، شبکه MLP برای اولین بار توسط IBBO آموزش داده می‌شود و نتایج با BBO استاندارد و الگوریتم‌های معیار استاندارد دیگر مقایسه خواهد شد.

$$\begin{cases} \mu_k = E \times \ln\left(\frac{k}{N} + 1\right) \\ \lambda_k = I \times \exp\left(-\frac{k}{N}\right) \end{cases} \quad k \leq N/p \quad (9)$$

$$\begin{cases} \mu_k = E \times \exp\left(\frac{k}{N} - 1\right) \\ \lambda_k = I \times \ln\left(2 - \frac{k}{N}\right) \end{cases} \quad k > N/p \quad (10)$$

که در این روابط، N نشان‌دهنده جمعیت اولیه است و



شکل (۳): شبکه‌ی عصبی چندلایه با ساختار ۲-۳-۱.

بعد از نمایش شبکه‌های عصبی چندلایه به صورت بردارهای محل سکونت، برای ارزیابی هر کدام از آن‌ها (محل‌های سکونت)، لازم است که رابطه‌ای برای محاسبه شاخص مناسب بودن محل سکونت (تابعی مناسب) نوشته شود.

۴-۲- شاخص مناسب بودن زیستگاه (تابع شایستگی)

همان‌طور که بیان شد، هدف نهایی روش‌های یادگیری، آموزش شبکه‌های عصبی مصنوعی است. مهم‌ترین بخش در یادگیری، فرآیند آموزش است. هر نمونه آموزش باید شامل، محاسبه شاخص مناسب بودن همه محل‌های سکونت باشد. در این مقاله، تابع شاخص مناسب بودن زیستگاه (برای همه نمونه‌های آموزش)، به روش میانگین مربعات خطا^۱ و به صورت رابطه (۱۲) محاسبه می‌گردد:

$$E = \sum_{k=1}^q \frac{\sum_{i=1}^m (o_i^k - d_i^k)^2}{q} \quad (12)$$

که در آن، q مبین تعداد نمونه‌های آموزش، m مبین تعداد خروجی‌ها، d_i^k مبین خروجی مطلوب از i -امین ورودی است وقتی که k -امین نمونه آموزش استفاده گردد و o_i^k مبین خروجی واقعی در ازای i -امین ورودی است وقتی که k -امین نمونه آموزش به ورودی اعمال می‌گردد. برای مثال، مقدار شاخص مناسب بودن محل سکونت برای i -امین محل سکونت به صورت رابطه (۱۳) محاسبه می‌گردد:

$$HSI(Habitat_i) = E(Habitat_i) \quad (13)$$

می‌توان در دو مرحله با استفاده از الگوریتم بهینه‌سازی مبتنی بر جغرافیای زیستی، یادگیری در شبکه عصبی چندلایه را فرمول نویسی نمود. بلوک دیاگرام پیشنهاد شده در شکل (۴) نشان داده شده است.

همان‌طوری که در شکل (۴) مشاهده می‌گردد، روش پیشنهاد شده با تولید مجموعه‌های تصادفی از شبکه‌های عصبی

(الف) راهبرد نمایش: در الگوریتم بهینه‌سازی مبتنی بر جغرافیای زیستی، باید وزن یال‌ها و بایاس گره‌ها به صورت مناسب نمایش داده شوند.

(ب) شاخص مناسب بودن محل سکونت: برای ارزیابی زیستگاه‌ها، باید یک تابع مناسب خطای شبکه عصبی چندلایه تعریف گردد. در بخش‌های بعدی این مراحل با جزئیات تشریح خواهند شد.

۴-۱- نمایش مسأله آموزش شبکه‌ی عصبی چندلایه توسط الگوریتم بهینه‌سازی مبتنی بر جغرافیای زیستی

به‌طور کلی، سه روش برای نمایش وزن یال‌ها و بایاس گره‌ها وجود دارد: بردار، ماتریس و حالت دودویی [۳۵]. در نمایش برداری، هر پارامتر با یک بردار نمایش داده می‌شود. برای آموزش یک شبکه عصبی چندلایه، باید همه وزن‌ها و بایاس‌ها معلوم باشند. در نمایش ماتریسی، هر پارامتر به صورت یک ماتریس نمایش داده می‌شود. برای نمایش دودویی، هر پارامتر به صورت رشته‌ای از بیت‌های دودویی نمایش داده می‌شود. هر کدام از این شیوه‌های نمایش، مزایا و معایب خاص خود را دارد که می‌تواند در یک مسأله خاص مفید واقع گردد [۳۶].

در روش نخست، تبدیل پارامترها به بردار، ماتریس و یا رشته‌ای از بیت‌های دودویی آسان می‌باشد، اما فرآیند بازیابی آن‌ها پیچیده خواهد بود. به همین دلیل، اغلب این روش در شبکه‌های عصبی ساده مورد استفاده قرار می‌گیرد. در روش دوم برای شبکه‌های با ساختار پیچیده، بازیابی آسان‌تر از کد کردن پارامترها است. این روش برای الگوریتم‌های یادگیری در شبکه‌های عصبی عمومی بسیار مناسب است. در روش سوم، نیاز است که متغیرها به صورت دودویی نمایش داده شوند. در این حالت وقتی که ساختار شبکه پیچیده گردد، طول هر پارامتر نیز افزایش می‌یابد. بنابراین، فرآیند کد کردن و دی‌کد کردن خیلی پیچیده خواهد شد.

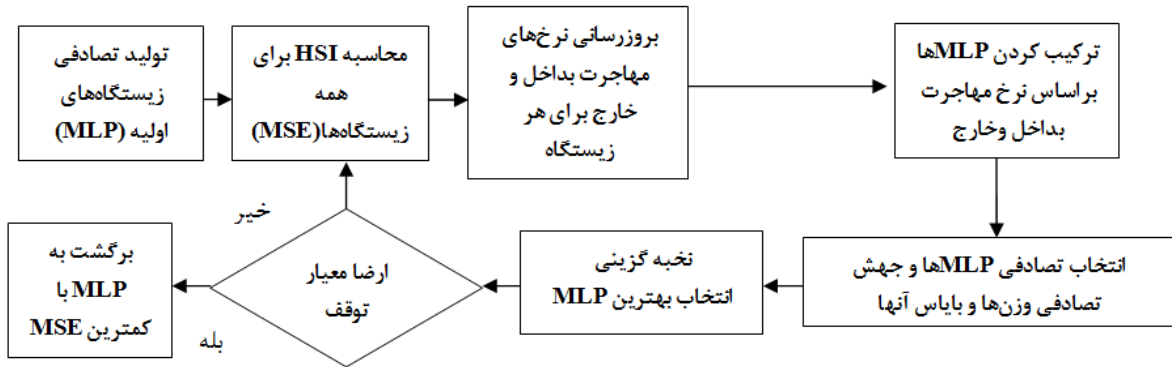
در این مقاله، چون با شبکه‌های عصبی چندلایه‌ی پیچیده سروکار نداریم، از روش برداری استفاده شده است. به‌منظور کاهش زمان اجرای برنامه شبکه‌های عصبی چندلایه، از جعبه ابزارهای عمومی Matlab استفاده نخواهد شد. به‌عنوان مثالی از این شیوه کدنویسی، بردار نهایی شبکه عصبی چندلایه نشان داده شده در شکل (۳)، در رابطه (۱۱) آورده شده است.

محل سکونت:

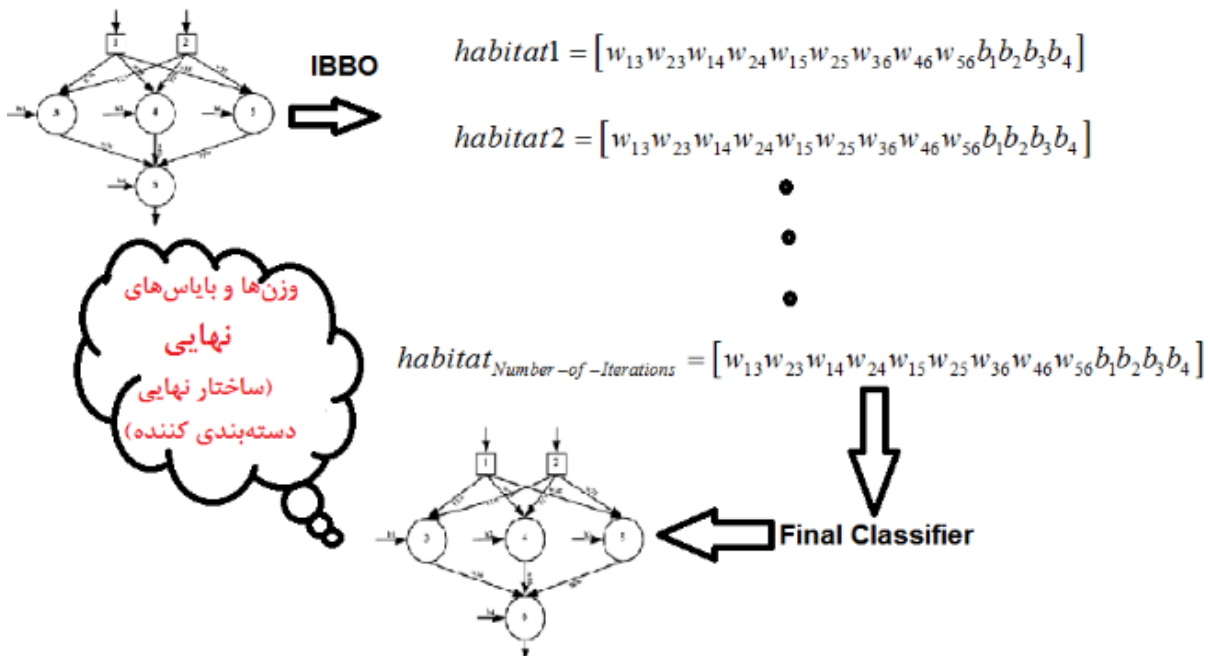
$$habitat = [w_{13} w_{23} w_{14} w_{24} w_{15} w_{25} w_{36} w_{46} w_{56} b_1 b_2 b_3 b_4] \quad (11)$$

چندلایه، بر مبنای نرخ جهش محل سکونتش تغییر می‌کند. انتخاب نخبه گام آخر روش پیشنهاد شده است، به طوری که بهترین شبکه‌های عصبی چندلایه به منظور جلوگیری از خرابی توسط عملگرهای جهشی و تکاملی در نسل بعدی، حفظ گردند. این مراحل (از محاسبه‌ی میانگین مربعات خطا تا انتخاب نخبه) تا ارضای شرایط پایانی ادامه می‌یابد. خلاصه مطالب بالا را می‌توان در شکل (۵) مشاهده کرد.

چندلایه بر مبنای تعداد محل‌های سکونت تعریف شده، آغاز می‌گردد. هر شبکه عصبی چندلایه متناظر با یک زیستگاه می‌باشد و هر وزن یا بایاس، با ساکنین آن زیستگاه متناظر است. بعد از گام آغازین، با استفاده از رابطه (۱۲) میانگین مربعات خطای هر شبکه‌های عصبی چندلایه محاسبه می‌گردد. در گام بعدی با استفاده از روابط (۶-۵)، نرخ‌های مهاجرت به بیرون، مهاجرت به داخل و جهش به‌روز می‌شوند. سپس شبکه‌های عصبی چندلایه براساس نرخ‌های مهاجرت به داخل و مهاجرت به بیرون با هم ترکیب می‌شوند. پس از آن هر شبکه عصبی



شکل (۴): بلوک‌دیگرام روش پیشنهاد شده.

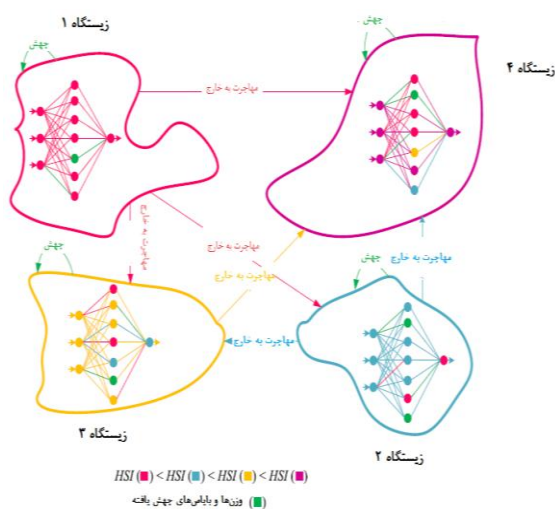


شکل (۵): خلاصه روش پیشنهاد شده.

یک شبکه عصبی چندلایه با h گره پنهان، o خروجی و تعداد t نمونه آموزش، برابر است با $O(t(h+o))$. در پیاده‌سازی انجام‌گرفته پیچیدگی محاسباتی مهاجرت برابر با $O(mn^2)$ است که در آن، m مبین تعداد ساکنین و n مبین تعداد محل‌های

پیچیدگی محاسبات روش پیشنهاد شده به تعداد نمونه‌های آموزش گیرنده در مجموعه داده‌ها، ساختار شبکه عصبی چندلایه، تعداد محل‌های سکونت، سازوکار مهاجرت، مکانیزم جهش و مکانیزم نخبه‌گزینی وابسته است. پیچیدگی محاسباتی

- قدرت اکتشاف بالاتر، الگوریتم بهینه‌سازی مبتنی بر جغرافیای زیستی را از گیرافتادن در حلقه‌های بهینه‌سازی محلی حفظ کرده و باعث رفع شرایط رکود (حالت ایستایی) در الگوریتم می‌گردد.
- از آن جایی که وزن‌ها و بایاس‌های شبکه‌های عصبی چندلایه بهتر تمایل به مهاجرت به سمت شبکه‌های بدتر دارند، در نتیجه میانگین‌های مربعات خطای (شاخص مناسب بودن محل سکونت) همه شبکه‌های عصبی چندلایه، بهبود می‌یابد و همین امر، هم‌گرایی روش پیشنهاد شده را تضمین می‌نماید و همه شبکه‌های عصبی چندلایه را بهبود می‌بخشد.
- نرخ‌های جهش متفاوت هر محل سکونت به روش پیشنهاد شده کمک می‌کند تا مکانیزم‌های بهره‌برداری متنوعی داشته باشد.
- نخبه‌گزینی (انتخاب بهترین محل سکونت) به روش پیشنهاد شده کمک می‌کند، به طوری که این راه‌حل‌ها هرگز از بین نروند.



شکل (۶): مهاجرت بین محل‌های سکونت برای یادگیری یک شبکه MLP [۱].

۵- تشریح مطالب و نتایج

در این بخش الگوریتم بهینه‌سازی مبتنی بر جغرافیای زیستی روی دادگان سونار با ابعاد کاهش یافته، Iris و Lenses اعمال می‌گردد [۳۸]. هم‌چنین به منظور اطمینان از صحت عملکرد الگوریتم IBBO، مجموعه دادگان با الگوریتم‌های بهینه‌سازی ازدحام ذرات، ژنتیک، بهینه‌سازی کلونی مورچگان و جستجوی گرانشی مقایسه خواهند شد.

۵-۱- تنظیم پارامترها و انجام آزمایش

پارامترهای مورد نیاز و مقادیر اولیه در جدول (۱) آورده شده‌اند.

سکونت است. ذکر این نکته لازم است که در بهترین حالت، پیچیدگی عملگر مهاجرت برابر است با $O(mn)$. پیچیدگی محاسباتی عملگر جهش در بدترین حالت برابر با $O(mn)$ است.

از آن جایی که در مرحله انتخاب نخبه، بهترین محل سکونت به روش مرتب‌کردن سریع^۱ انتخاب می‌گردد، در بهترین حالت، پیچیدگی محاسباتی نخبه‌گزینی (انتخاب بهترین زیستگاه) برابر با $O(n \cdot \log(n))$ و در بدترین حالت برابر با $O(n^2)$ خواهد بود. بنابراین پیچیدگی محاسباتی نهایی روش پیشنهادی به صورت رابطه (۱۴) محاسبه می‌گردد:

$$O(MLP, BBO) = O\left(g\left(t(h+o) + mn^2 + nm + n^2\right)\right) \quad (14)$$

که در آن، g مبین بیش‌ترین تعداد نسل‌ها، t مبین تعداد نمونه‌های آموزش، h مبین تعداد گره‌های پنهان، o مبین تعداد گره‌های خروجی، m مبین تعداد ساکنین و n مبین تعداد زیستگاه‌ها می‌باشد.

برای مشاهده چگونگی کار الگوریتم پیشنهادی، در شکل (۶) یک تصویر مفهومی از مهاجرت بین محل‌های سکونت برای یادگیری در یک شبکه عصبی چندلایه با استفاده از الگوریتم بهینه‌سازی مبتنی بر جغرافیای زیستی، نشان داده شده است. در این شکل، محل سکونت ۱ مناسب‌تر از محل‌های سکونت ۲، ۳ و ۴ می‌باشد، زیرا کم‌ترین شاخص مناسب بودن محل سکونت را دارد که نشان می‌دهد میانگین مربعات خطا برای همه نمونه‌های آزمایش، حداقل است.

همان‌طور که ملاحظه می‌شود، زیستگاه ۱ بیش‌ترین مهاجرت به بیرون را دارد، در حالی که زیستگاه ۴ بیش‌ترین مهاجرت به داخل را دارد و در نتیجه ساکنین بیش‌تری (وزن‌ها و بایاس‌ها) را نسبت به محل‌های سکونت دیگر می‌پذیرد. این مهاجرت‌ها با رنگ‌های متفاوتی نشان داده شده‌اند. هم‌چنین گره‌های سبز رنگ و اتصالات آن‌ها جهش را نشان می‌دهند که برای همه ساکنین بدون در نظر گرفتن مقدار شاخص مناسب بودن محل سکونت اتفاق افتاده است. برای این که ببینید به صورت نظری چگونه الگوریتم بهینه‌سازی مبتنی بر جغرافیای زیستی، آموزش شبکه‌های عصبی چندلایه را بهبود می‌بخشد، مشاهدات زیر را در نظر بگیرید:

- مقادیر مختلف نرخ مهاجرت به بیرون و مهاجرت به داخل باعث ایجاد مکانیزم تکاملی برای هر محل سکونت می‌گردد که قدرت اکتشاف بیش‌تر را در پی دارد.

می‌شود. نرخ دسته‌بندی و درصد خطای آزمون دو مقیاس برای مقایسه الگوریتم‌های ذکر شده می‌باشند. برای انجام یک مقایسه نسبتاً خوب، همه الگوریتم‌ها وقتی که حداکثر تعداد تکرار به ۲۵۰ رسید، متوقف می‌شوند. در نهایت، هم‌گرایی نتایج برای انجام مقایسه‌ای جامع بررسی خواهد شد. در دسته‌بندی مجموعه‌های داده، از آن جایی که برای انتخاب تعداد گره‌های پنهان استاندارد وجود ندارد. بنابراین، براساس ساختار شبکه‌های عصبی چندلایه، از پیشنهاد مطرح شده در [۳۷] و از رابطه (۱۵) استفاده خواهد شد.

$$H = 2 \times N + 1 \quad (15)$$

که در آن، N مبین تعداد ورودی‌ها و H مبین تعداد گره‌های پنهان است.

در ابتدا دسته‌بندی‌کننده طراحی شده بر روی دادگان Iris و Lenses (توصیف شده در جدول (۲)) اعمال شده و عملکرد دسته‌بندی‌کننده از نظر نرخ دسته‌بندی، اجتناب از گیرکردن در کمینه محلی و سرعت هم‌گرایی آزموده می‌شود. هر الگوریتم ۱۰ بار اجرا شده و نرخ دسته‌بندی، میانگین و انحراف معیار حداقل خطا و مقدار P-value در جداول (۳-۴) به ترتیب برای مجموعه دادگان Iris و Lenses نمایش داده شده است. نرخ دسته‌بندی، دقت دسته‌بندی‌کننده طراحی شده را نشان می‌دهد و مقادیر میانگین و انحراف معیار حداقل خطا و P-value نشان‌دهنده قدرت الگوریتم در اجتناب از بهینه محلی می‌باشد [۳]. نمونه‌ای از نتایج این مقایسه به ترتیب برای دادگان Iris و Lenses در اشکال (۸-۷) نمایش داده شده است. سپس در قسمت بعد دادگان سونار به‌طور مفصل توضیح داده شده و دسته‌بندی‌کننده‌های طراحی شده را بر روی این نوع دادگان نیز آزمایش می‌کنیم.

جدول (۲): دادگان مورد استفاده در این مقاله.

نام	وظیفه پیش فرض	ویژگی مشخصه	تعداد ویژگی	تعداد مثال	سال
Iris	دسته‌بندی	چندمتغیری	۴	۱۵۰	۱۹۸۸
Lenses	دسته‌بندی	چندمتغیری	۴	۲۴	۱۹۹۰

جدول (۱): پارامترهای اولیه الگوریتم‌ها.

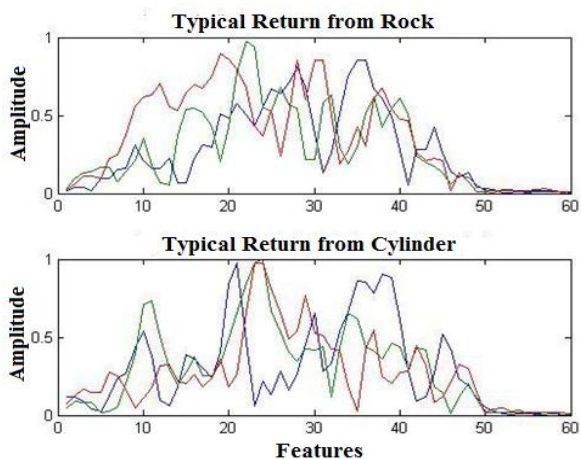
الگوریتم	پارامتر	مقدار
BBO و IBBO	احتمال اصلاح ساکنین	۱
	محدوده احتمال مهاجرت به‌داخل	[۰, ۱]
	اندازه پله برای انتگرال عددی احتمال	۱
	حداکثر ضریب مهاجرت به‌داخل (I) و مهاجرت به خارج (E)	۱
	احتمال جهش	۰/۰۰۵
	اندازه جمعیت	۲۰۸
	حداکثر تعداد تکرار	۲۵۰
PSO	جانمایی	اتصال کامل
	ثابت شناختی (C_1)	۱
	ثابت اجتماعی (C_2)	۱
	ثابت محلی (W)	۰/۳
GA	اندازه جمعیت	۲۰۸
	نوع	کدشده واقعی
	انتخاب	چرخ رولت
	ادغام	تک نقطه‌ای
	جهش	یکنواخت (۰/۰۱)
	اندازه جمعیت	۲۰۸
	حداکثر تعداد تکرار	۲۵۰
ACO	فرمون اولیه (T_0)	۰/۰۰۰۰۰۱
	ثابت به‌روزرسانی فرمون (Q)	۲۰
	ثابت فرمون (q_0)	۱
	نرخ کاهش فرمون سراری (p_g)	۰/۹
	نرخ کاهش فرمون محلی (p_l)	۰/۵
	حساسیت فرمون (α)	۱
	حساسیت قابل رویت	۵
GSA	اندازه جمعیت	۲۰۸
	تعداد جرم‌ها	۶۰
	G_0	۱
	α	۲۰
	حداکثر تکرار	۵۰۰

هر شبکه ۱۰ بار آزمایش شده است. بهترین شبکه عصبی آموزش دیده از بین ۱۰ بار اجرا، انتخاب و برای مقایسه به‌کار گرفته

۵-۲- دادگان سونار

دادگان مورد استفاده در این مقاله از آزمایش Gorman و Sejnowski موجود در مراجع [۳۸-۳۹] استخراج شده است. در این آزمایش یک سیلندر فلزی به طول ۵ فوت و یک صخره هم‌اندازه با آن در بستر شنی دریا قرار داده شده‌اند و یک پالس چیرپ FM خطی پهن باند (ka=55/6) به سمت آن‌ها فرستاده شده است. اکوهای برگشتی در فاصله ۱۰ متری از آن‌ها جمع‌آوری شده است.

بر اساس SNR اکوی دریافتی از ۱۲۰۰ اکو، ۲۰۸ اکو که SNR آن‌ها بین ۴ dB تا ۱۵ dB است، انتخاب شده‌اند. از این ۲۰۸ اکو، ۱۱۱ عدد مربوط به سیلندر فلزی و ۹۷ عدد مربوط به صخره هستند. شکل (۹) نمونه‌هایی از اکوهای دریافتی از صخره و سیلندر فلزی را نشان می‌دهد.



شکل (۹): نمایش دامنه اکوهای برگشتی از سیلندر فلزی و صخره.

از نقطه نظر ریاضی، PCA سعی در پیدا کردن یک نگاهت خطی به نام M می‌نماید که تابع هزینه رابطه (۱۶) را حداکثر نماید.

۵-۳- کاهش ابعاد دادگان با استفاده از روش تحلیل عناصر اصلی

تحلیل عناصر اصلی (PCA) یک روش خطی برای کاهش ابعاد دادگان می‌باشد. این روش به وسیله تعبیه دادگان در زیرفضاهای خطی با ابعاد پایین‌تر عمل می‌کند. در واقع این روش محورهای مختصات را به گونه‌ای می‌چرخاند که دادگان با حداکثر واریانس در راستای محورهای اصلی قرار بگیرند. اگرچه روش‌های زیادی برای کاهش ابعاد در سال‌های اخیر معرفی شده است، ما در این مقاله از روش PCA بدون ناظر و کلاسیک استفاده کرده‌ایم. همان‌گونه که در مرجع [۴۰] اثبات شده است، روش‌های جدید و ترکیبی برای دادگان مجازی دارای عملکرد بسیار بهتری نسبت

جدول (۳): نتایج حاصل از اعمال دسته‌بندی‌کننده‌های مختلف بر روی

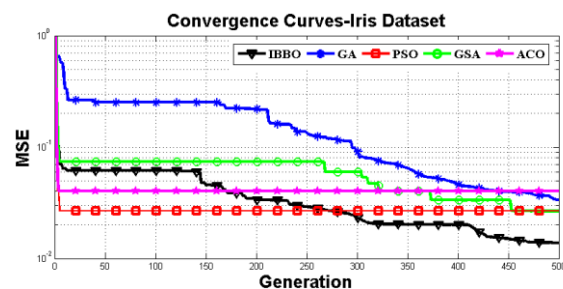
دادگان Iris.

الگوریتم	نرخ دسته‌بندی	P-Values	MSE (AVE±STD)
MLPIBBO	۸۸/۰۶۶۶	N/A	۰/۰۰۱۴±۰/۰۰۱۴۵
MLPGA	۸۵/۳۵۲۲٪	۰/۰۰۱۱	۰/۱۲۲۵±۰/۱۲۴۷
MLPPSO	۸۷/۸۲۲۲٪	۰/۰۰۰۴۷	۰/۱۳۱۱±۰/۱۱۴۷
MLACO	۸۳/۶۳۳۳٪	۰/۰۰۰۱۷	۰/۱۷۲۹±۰/۲۳۳۳
MLPGSA	۸۵/۳۳۳۳٪	۰/۰۰۰۱۸	۰/۱۵۷۷±۰/۲۲۲۴

جدول (۴): نتایج حاصل از اعمال دسته‌بندی‌کننده‌های مختلف بر روی

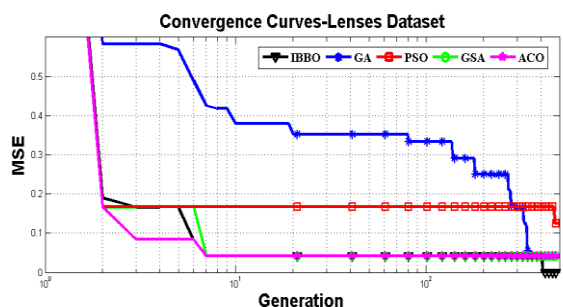
دادگان Lenses.

الگوریتم	نرخ دسته‌بندی	P-Values	MSE (AVE±STD)
MLPIBBO	۸۹/۸۹۳۳٪	N/A	۰/۱۲۸۳±۰/۰۰۱۴
MLPGA	۸۴/۲۲۲۲٪	۰/۰۰۴۷	۰/۱۵۱۹±۰/۰۲۶۹
MLPPSO	۸۲/۶۶۶۶٪	۰/۰۰۷۸	۰/۲۰۱۱±۰/۲۰۷۶
MLACO	۷۲/۲۲۳۳٪	۰/۰۰۰۷	۰/۳۱۴۹±۰/۲۹۶۵
MLPGSA	۷۳/۶۶۶۶٪	۰/۰۰۰۷	۰/۲۵۲۷±۰/۱۷۴۴



شکل (۷): مقایسه دقت دسته‌بندی و هم‌گرایی الگوریتم‌های مختلف

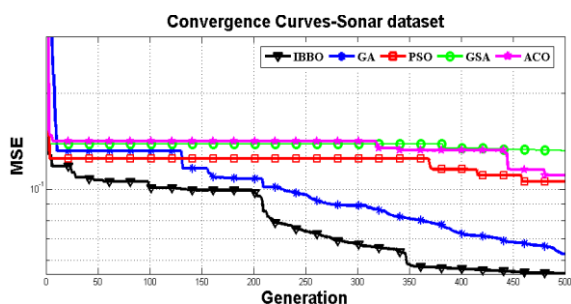
اعمال شده به دادگان Iris.



شکل (۸): مقایسه دقت دسته‌بندی و هم‌گرایی الگوریتم‌های مختلف

اعمال شده به دادگان Lenses.

از نقطه‌نظر آماری، الگوریتم بهینه‌سازی مبتنی بر جغرافیای زیستی، به‌اندازه کافی توانایی جلوگیری از گرفتادن در حداقل‌های محلی را دارد. کارایی بهتر الگوریتم IBBO در جستجوی همه فضای مسأله می‌باشد. به‌دلیل ماهیت طبیعی این الگوریتم‌ها و به‌خاطر مهاجرت و جهش در جلوگیری از گرفتادن در حداقل‌های محلی این عمل به خوبی انجام می‌پذیرد. این عملگرها باعث تغییرات ناگهانی در انتخاب راه‌حل مسأله می‌گردند که در نتیجه توانایی شناسایی الگوریتم‌های بهینه‌سازی مبتنی بر جغرافیای زیستی و ژنتیک را به‌اندازه کافی بهبود می‌دهد.



شکل (۱۱): مقایسه دقت دسته‌بندی و هم‌گرایی الگوریتم‌های مختلف اعمال‌شده به دادگان سونار.

جدول (۵): نتایج حاصل از اعمال دسته‌بندی‌کننده‌های مختلف بر روی دادگان سونار

الگوریتم	نرخ دسته‌بندی	P-Values	MSE (AVE±STD)
MLPBBO	۸۸/۲۳۵٪	۰/۰۰۳۹	۰/۱۲۱۹±۰/۰۲۶۹
MLPGA	۶۵/۸۶۲٪	۰/۰۰۱۷	۰/۳۵۱۱±۰/۱۰۷۶
MLPPSO	۸۷/۳۳۳٪	۰/۰۰۰۱	۰/۱۳۴۵±۰/۰۹۶۵
MLACO	۷۳/۹۳۳٪	۰	۰/۱۷۲۲±۰/۱۰۶۴
MLPGSA	۸۰/۱۱۱۲٪	۰	۰/۲۰۱۹±۰/۱۴۲۳

نتایج ضعیف الگوریتم‌های بهینه‌سازی کلونی مورچگان و انبوه ذرات نیز به‌دلیل ماهیت طبیعی این الگوریتم‌ها است. این الگوریتم‌ها عملگری را برای تغییر ناگهانی راه‌حل مسأله ندارند و در نتیجه در حداقل‌های محلی گرفتار می‌شوند. علاوه براین، الگوریتم بهینه‌سازی کلونی مورچگان، از ماتریس فرمونی که قدرت یادگیری و بهره‌برداری الگوریتم را افزایش می‌دهد، استفاده می‌کند که یک مزیت در مسائل ترکیبی است، ولی احتمال گرفتادن در حداقل‌های محلی را افزایش می‌دهد. الگوریتم‌های انبوه ذرات به‌مقدار خیلی زیاد به نحوه توزیع اولیه ذرات و محرک‌های اولیه آن‌ها بر مبنای جاذبه بینشان، وابسته است. اگر تعداد زیادی از ذرات در حداقل‌های محلی گرفتار شوند، الگوریتم به‌مقدار کوچکی از گرفتادن ذرات دیگر جلوگیری خواهد کرد.

به روش‌های دسته یک هستند، اما روش‌های دسته یک مانند PCA برای دادگان دنیای واقعی دارای عملکرد بسیار بهتری می‌باشند. از طرف دیگر، برای سنجش عملکرد PCA طراحی‌شده دو روش وجود دارد که عبارتند از: الف) سنجش همراه با ناظر و ب) سنجش بدون ناظر. روش با ناظر از نظر دقت بسیار کارآمد می‌باشد، این در حالی است که از لحاظ زمانی بسیار پرهزینه می‌باشد. به‌دلیل نقیصی که در روش اول وجود دارد (افزایش بار محاسباتی و پیچیدگی زمانی)، در این مقاله از روش دوم (بدون ناظر) استفاده کرده‌ایم.

$$trace(\mathbf{M}^T \text{cov}(\mathbf{X})\mathbf{M}) \quad (۱۶)$$

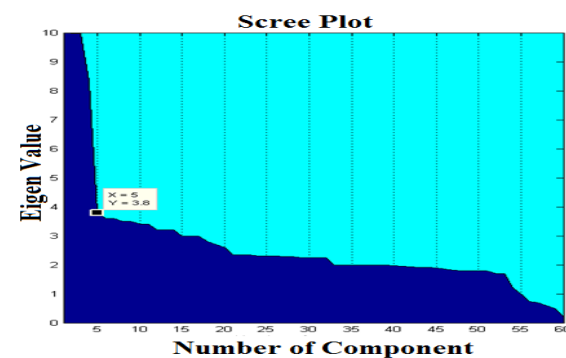
در این رابطه، $\text{cov}(\mathbf{X})$ ماتریس کواریانس دادگان \mathbf{X} می‌باشد. می‌توان نشان داد که این نگاشت خطی به‌وسیله d بردار ویژگی اصلی ماتریس کواریانس دادگان با میانگین صفر شکل می‌گیرد. از این رو این نوع PCA سعی در حل رابطه (۱۷) دارد.

$$\text{cov}(\mathbf{X})\mathbf{M} = \lambda\mathbf{M} \quad (۱۷)$$

با حل این رابطه می‌توان d مقدار ویژه λ را به‌دست آورد. دادگان اصلی x_i با ابعاد D (در این مقاله $D = 60$) توسط نگاشت \mathbf{M} و رابطه (۱۸) به دادگان جدید y_i با ابعاد d که $d < D$ است، تبدیل می‌شوند.

$$\mathbf{Y} = \mathbf{X}\mathbf{M} \quad (۱۸)$$

انتخاب مقدار d یکی دیگر از مسائل چالش‌برانگیز این روش می‌باشد. در این مقاله برای انتخاب d از روش Scree Test استفاده شده است. این روش [۴۱] برای انتخاب d از نمایش نزولی نمودار مقادیر ویژه استفاده می‌کند. بدین صورت که نقطه‌ای را که نمودار افت شدیدی دارد، ولی هنوز در مقادیر کوچک ثابت نشده است را به‌عنوان d انتخاب می‌کند. نتیجه این آزمایش در شکل (۱۰) نمایش داده شده است. همان‌گونه که در این شکل دیده می‌شود، برای دادگان سونار، $9 < d < 6$ می‌تواند انتخاب شود که برای دقت بیش‌تر $d = 9$ انتخاب شده است.



شکل (۱۰): نتایج حاصل از انجام Scree Test بر روی دادگان سونار

زیستی در آموزش شبکه‌های عصبی چندلایه، مورد استفاده قرار گرفت. نتایج آماری به دست آمده با نتایج حاصل از الگوریتم بهینه‌سازی گروهی ذرات، الگوریتم ژنتیک، الگوریتم کلونی مورچه‌ها و الگوریتم جستجوی گرانشی، جهت اثبات کارایی، مقایسه گردید. نتایج حاصل نشان می‌دهد که الگوریتم بهینه‌سازی مبتنی بر جغرافیای زیستی به اندازه کافی توانایی جلوگیری از گیرافتادن در حداقل‌های محلی را در مقابل الگوریتم‌های معیار دارد. به علاوه، به وضوح کارایی بهتر الگوریتم IBBO در آموزش شبکه‌های عصبی چندلایه بر حسب دقت نتایج و سرعت هم‌گرایی از نتایج به دست آمده دیده می‌شود.

۷- مراجع

- [1] M. R. Mosavi, M. Khishe, and M. Aghababaie, "Modeling and Mitigation of Active Sonar Clutter," Noshahr University of Marine Science and Technology, 2015.
- [2] M. R. Mosavi, M. Khishe, A. Ghamgosar and M. J. Ghalandari, "Classification of Sonar Data Set using the Gray Wolf Optimizer Algorithm", Journal of Electronics Industries, Vol.7, No.1, pp.27-41, 1395. (In Persian)
- [3] M. R. Mosavi, M. Khishe, and E. Ebrahimi, "Classification of Sonar Targets using OMKC," Genetic Algorithm and Statistical Moments," Journal of Advances in Computer Research, vol.7, no.1, pp. 143-156, 2016.
- [4] L. S. Nguyen, D. Frauendorfer, M. S. Mast, and D. Gatica-Perez, "Hire Me: Computational Inference of Hirability in Employment Interviews based on Nonverbal Behavior," IEEE Transactions on Multimedia, vol. 16, no. 4, pp. 1018-1031, 2014.
- [5] M. R. Mosavi, M. Khishe, Y. Hatam Khani, and M. Shabani, "Training Radial Basis Function Neural Network using Stochastic Fractal Search Algorithm to Classify Sonar Dataset," Iranian Journal of Electrical and Electronic Engineering, vol. 13, no. 1, 2017.
- [6] E. Oja, "Unsupervised Learning in Neural Computation," Theoretical Computer Science, vol. 287, pp. 187-207, 2002.
- [7] N. Zhang, "An Online Gradient Method with Momentum for Two-Layer Feedforward Neural Networks," Applied Mathematics and Computation, vol. 212, pp. 488-498, 2009.
- [8] D. R. Hush and B. G. Horne, "Progress in Supervised Neural Networks," IEEE Signal Processing Magazine, vol. 10, pp. 8-39, 1993.
- [9] S. C. Ng, C. C. Cheung, S. H. Leung, and A. Luk, "Fast Convergence for Backpropagation Network with Magnified Gradient Function," IEEE Joint Conference on Neural Networks, vol. 3, pp. 1903-1908, 2003.
- [10] G. Magoulas, M. Vrahatis, and G. Androulakis, "On the Alleviation of the Problem of Local Minima in Back-Propagation," Nonlinear Analysis, Theory, Methods & Applications, vol. 30, no. 7, pp. 4545-4550, 1997.
- [11] Y. C. Ho and D. L. Pepyne, "Simple Explanation of the No-Free-Lunch Theorem and Its Implications," Journal of Optimization Theory and Applications, vol. 115, no. 3, pp. 549-570, 2002.
- [12] P. Wang, X. Yu, and J. Lu, "Identification and Evolution of Structurally Dominant Nodes in Protein-Protein Interaction

دلیل کارایی بهتر الگوریتم بهینه‌سازی مبتنی بر جغرافیای زیستی در مقایسه با الگوریتم ژنتیک در اکثر مسائل، نرخ‌های متفاوت مهاجرت به بیرون و مهاجرت به داخل هر محل سکونت است. الگوریتم بهینه‌سازی مبتنی بر جغرافیای زیستی در مقایسه با الگوریتم ژنتیک که برای همه محل‌های سکونت جامعه آماری آن، یک نرخ تولید مجدد کلی دارد، دو نرخ (مهاجرت به بیرون و مهاجرت به داخل) برای هر زیستگاه دارد. این مسئله باعث ایجاد رفتار تکاملی و قدرت شناسایی متفاوت می‌گردد. به طور خلاصه، می‌توانیم بگوییم در مسئله آموزش شبکه‌های عصبی چندلایه، توانایی اکتشاف، بسیار مهم است. بنابراین، در حل مسائل پیچیده با استفاده از شبکه‌های عصبی چندلایه، به گام‌های جستجوی تصادفی و ناگهانی برای جلوگیری از گیرافتادن در حداقل‌های محلی نیاز است.

این مقاله نشان داد که عملگرهای (مهاجرت) الگوریتم بهینه‌سازی مبتنی بر جغرافیای زیستی، برای این منظور بسیار مناسب هستند. در ادامه دلایل برتری کارایی IBBO در مقایسه با الگوریتم پس‌انتشار و سایر الگوریتم‌های تکاملی آورده شده است:

- مقادیر متفاوت نرخ‌های مهاجرت به بیرون و داخل اطلاعات گوناگونی در خصوص تغییر محل‌های سکونت و در نتیجه آن، بهبود اکتشاف را در پی دارد.
- در مدت تولید، از آنجایی که ساکنین در محل‌های سکونت با شاخص بالا تمایل به مهاجرت به محل‌های سکونت با شاخص پایین دارند، شاخص مناسب بودن همه زیستگاه‌ها بهبود می‌یابد. این امر، هم‌گرایی الگوریتم بهینه‌سازی مبتنی بر جغرافیای زیستی را تضمین می‌کند.
- عملگرهای مهاجرت توانایی اکتشاف الگوریتم را افزایش می‌دهند و در نتیجه الگوریتم بهینه‌سازی مبتنی بر جغرافیای زیستی در حداقل‌های محلی گرفتار نمی‌شود.
- نرخ‌های جهش متفاوت باعث متنوع شدن محل‌های سکونت می‌گردد.
- نخبه‌گزینی به الگوریتم‌های بهینه‌سازی مبتنی بر جغرافیای زیستی کمک می‌کند تا بتوان راه‌حل‌های مناسب را ذخیره و دوباره مورد استفاده قرار داد و بنابراین هیچ‌گاه از روند محاسبات کنار نروند.

۶- نتیجه‌گیری

در این مقاله، یک شبکه عصبی چندلایه با استفاده از الگوریتم IBBO، آموزش داده شد. دادگان سونار با ابعاد کاهش‌یافته، به منظور بررسی تأثیر الگوریتم بهینه‌سازی مبتنی بر جغرافیای

- V. Abedifar, M. Eshghi, S. Mirjalili, and S. M. Mirjalili, "An Optimized Virtual Network Mapping using PSO in Cloud Computing," 21st Iranian Conference on Electrical Engineering, pp. 1-6, 2013.
- [29] M. R. Mosavi, M. Kaveh, and M. Khishe, "Sonar Data Set Classification using MLP Neural Network Trained by Non-linear Migration Rates BBO," The Fourth Iranian Conference on Engineering Electromagnetic (ICEEM 2016), pp. 1-5, March 2016.
- [30] W. Guo, L. Wang, and Q. Wu, "An Analysis of the Migration Rates for Biogeography-based Optimization," *Information Science*, vol. 254, pp. 111-140, 2014.
- [31] M. R. Mosavi, M. Kaveh, M. Khishe, and M. Aghababae, "Design and implementation a Sonar Data Set Classifier by using MLP NN Trained by Improved Biogeography-based Optimization," The Second National Conference on Marine Technology, MMT, 2016.
- [32] M. Ergezer, D. Simon, and D. Du, "Oppositional Biogeography-based Optimization", IEEE Conference on Systems, Man and Cybernetics, San Antonio, Texas, pp. 1035-1040, 2009.
- [33] R. Rarick, D. Simon, F. E. Villaseca, and B. Vyakaranam, "Biogeography-based Optimization and the Solution of the Power Flow Problem," IEEE Conference on Systems, Man and Cybernetics, San Antonio, Texas, pp. 1029-1034, 2009.
- [34] J. R. Zhang, J. Zhang, T. M. Lok, and M. R. Lyu, "A Hybrid Particle Swarm Optimization-Back-propagation Algorithm for Feedforward Neural Network Training," *Applied Mathematics and Computation*, vol. 185, pp. 1026-1037, 2007.
- [35] S. Mirjalili, S. Z. M. Hashim, and H. M. Sardroudi, "Training Feedforward Neural Networks using Hybrid Particle Swarm Optimization and Gravitational Search Algorithm," *Applied Mathematics and Computation*, vol. 218, pp. 11125-11137, 2012.
- [36] S. Mirjalili, "Hybrid Particle Swarm Optimization and Gravitational Search Algorithm for Multilayer Perceptron Learning," *University Teknologi Malaysia (UTM)*, 2011.
- [37] [http://archive.ics.uci.edu/ml/datasets/Connectionist+Bench+\(Sonar,+Mines+vs.+Rocks\)](http://archive.ics.uci.edu/ml/datasets/Connectionist+Bench+(Sonar,+Mines+vs.+Rocks)).
- [38] R. P. Gorman and T. J. Sejnowski, "Analysis of Hidden Units in a Layered Network Trained to Classify Sonar Targets," *Neural Networks*, vol. 1, pp. 75-89, 1988.
- [39] S. C. H. Hoi, R. Jin, P. Zhao, and T. Yang, "Online Multiple Kernel Classification," *Machine Learning*, vol. 90, no. 2, pp. 289-316, 2013.
- [40] R. K. Jade, L. K. Verma, and K. Verma, "Classification using Neural Network and Support Vector Machine for Sonar Data Set," *International Journal of Computer Trends and Technology*, vol. 4, no. 2, pp. 116-119, 2013.
- Networks," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 8, no. 1, pp. 87-97, 2014.
- [13] V. G. Gudise and G. K. Venayagamoorthy, "Comparison of Particle Swarm Optimization and Backpropagation as Training Algorithms for Neural Networks," *IEEE Swarm Intelligence Symposium*, pp. 110-117, 2003.
- [14] R. Mendes, P. Cortez, M. Rocha, and J. Neves, "Particle Swarms for Feedforward Neural Network Training," *IEEE Joint Conference on Neural Networks*, pp. 1895-1899, 2002.
- [15] U. Seiffert, "Multiple Layer Perceptron Training using Genetic Algorithms," *European Symposium on Artificial Neural Networks*, pp. 159-164, 2001.
- [16] C. Blum and K. Socha, "Training Feed-forward Neural Networks with Ant Colony Optimization: An Application to Pattern Classification," *Hybrid Intelligent Systems Conference*, pp. 6-14, 2005.
- [17] G. Li, J. Na, D. Stoten, and X. Ren, "Adaptive Neural Network Feedforward Control for Dynamically Substructured Systems," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 3, pp. 944-954, 2014.
- [18] I. Boussaid, J. Lepagnot, and P. Siarry, "A Survey on Optimization Metaheuristics," *Information Sciences*, vol. 237, pp. 82-117, 2013.
- [19] S. Mirjalili and S. Z. M. Hashim, "A New Hybrid PSO-GSA Algorithm for Function Optimization," *IEEE Conference on Computer and Information Application*, pp. 374-377, 2010.
- [20] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," *Advances in Engineering Software*, vol. 69, pp. 46-61, 2014.
- [21] S. Mirjalili, S. M. Mirjalili, and X. S. Yang, "Binary Bat Algorithm," *Neural Computing and Applications*, vol. 25, no. 3-4, pp. 663-681, 2013.
- [22] M. R. Mosavi and M. Khishe, "Training a Feed-Forward Neural Network using Particle Swarm Optimizer with Autonomous Groups for Sonar Target Classification," *Journal of Circuits, Systems, and Computers (JCSC)*, vol. 26, no. 11, November 2017.
- [23] M. Khishe, M. R. Mosavi, and M. Kaveh, "Improved Migration Models of Biogeography-based Optimization for Sonar Data Set Classification using Neural Network," *Applied Acoustic*, vol. 118, pp. 15-29, 2017.
- [24] S. M. Mirjalili, S. Mirjalili, and A. Lewis, "A Novel Multi-Objective Optimization Framework for Designing Photonic Crystal Waveguides," *Photonics Technology Letters*, vol. 26, no. 2, pp. 146-149, 2014.
- [25] S. M. Mirjalili, S. Mirjalili, A. Lewis, and K. Abedi, "A Tri-Objective Particle Swarm Optimizer for Designing Line Defect Photonic Crystal Waveguides," *Photonics and Nanostructures Fundamentals and Applications*, vol. 12, no. 2, pp. 152-163, 2014.
- [26] S. Saremi, S. M. Mirjalili, and S. Mirjalili, "Unit Cell Topology Optimization of Line Defect Photonic Crystal Waveguide," *Procedia Technology*, vol. 12, pp. 174-179, 2014.
- [27] S. Saremi, S. M. Mirjalili, and S. Mirjalili, "Chaotic Krill Herd Optimization Algorithm," *Procedia Technology*, vol. 12, pp. 180-185, 2014.
- [28] M. R. Mosavi, M. Khishe, and M. Akbarisani, "Neural Network Trained by Biogeography-based Optimizer with Chaos for Sonar Data Set Classification," *Wireless Personal Communications (WPC)*, pp. 1-20, 2017.