

آموزش شبکه عصبی MLP در طبقه‌بندی داده‌ها با استفاده از روش GSA

مریم ده‌باشیان و سیدحمید ظهیری

شبکه‌های عصبی مصنوعی ساختارهای قدرتمندی برای فرایندهای کنترلی، طبقه‌بندی داده‌ها و خوشه‌بندی اطلاعات دارند. اوزان یک شبکه عصبی مصنوعی مرحله به مرحله بر اساس اطلاعات ورودی که دریافت می‌کند تغییر کرده و هر چه میزان این اطلاعات بیشتر شود شبکه خطای خود را بیشتر کاهش می‌دهد [۷].

یکی از مرسوم‌ترین انواع شبکه‌های عصبی، شبکه عصبی پرسپترون چندلایه^۴ (MLP) است که به‌طور موفقیت‌آمیزی در بازه وسیعی از کاربردها از جمله طبقه‌بندی داده مورد استفاده قرار گرفته است. هنگام کار با شبکه‌های عصبی MLP با دو مسئله روبه‌رو هستیم: انتخاب معماری مناسب و انتخاب الگوریتم آموزشی مناسب. معماری مناسب به معنی انتخاب بهینه تعداد لایه‌ها، تعداد نرون‌ها در هر لایه و نوع تابع تحریک هر نرون می‌باشد و معماری بهینه شبکه‌های عصبی مبتنی بر مجموعه داده‌ها و ویژگی‌های آنهاست. الگوریتم‌های آموزشی متنوعی جهت آموزش شبکه‌های عصبی به کار می‌رود. متداول‌ترین الگوریتم آموزشی این شبکه‌ها، الگوریتم پس‌انتشار خطا^۵ می‌باشد. در الگوریتم پس‌انتشار خطا در هر مرحله مقدار خروجی محاسبه‌شده جدید، با مقدار واقعی مقایسه شده و با توجه به خطای به‌دست آمده به اصلاح وزن‌های شبکه پرداخته می‌شود. به نحوی که در انتهای هر تکرار اندازه خطای حاصله کمتر از میزان به‌دست آمده در تکرار قبلی باشد. اساس این کمینه‌سازی، حرکت بر روی بردار گرادیان تابع مربعات خطای شبکه می‌باشد. مشکل الگوریتم پس‌انتشار خطا، همگرایی دیر و توقف در نقاط بهینه محلی است [۸].

تلاش‌های فراوانی برای سرعت‌بخشیدن به همگرایی و بهبود دقت الگوریتم پس‌انتشار خطا در آموزش شبکه‌های عصبی صورت گرفته است. از آن جمله می‌توان به استفاده از تطبیق مومنتوم [۹] و نرخ یادگیری متغیر [۱۰] اشاره کرد که به بهبود اندکی منجر شدند. همچنین با بزرگ کردن مصنوعی خطا برای نرون‌های عمل‌کننده در ناحیه اشباع نتایج بهتری به‌دست آمدند [۱۱]. با استفاده از روش‌های مختلف مرتبه دوم (برای مثال روش نیوتن، گرادیان مختلط و یا تکنیک بهینه‌سازی Levenberg - Marquardt [۱۲]) بهبود قابل توجهی را می‌توان بر روی عملکرد تشخیص مشاهده کرد، که به‌طور وسیعی به‌عنوان یکی از مؤثرترین روش‌ها در دقت تشخیص پذیرفته شده است.

در [۱۳] تا [۱۵] برای سرعت‌بخشیدن به فرایند آموزش، منطق فازی با الگوریتم‌های آموزش شبکه عصبی یکپارچه شده است. در [۱۶] نیز از روش بیز برای آموزش شبکه عصبی استفاده شده است.

رویکرد جدید در آموزش شبکه‌های عصبی استفاده از الگوریتم‌های ابتکاری است. روش‌های جستجوی ابتکاری الگوریتم‌هایی هستند که با الهام از فرایندهای فیزیکی و بیولوژیکی در طبیعت به‌وجود آمده‌اند و غالب

چکیده: تاکنون شیوه‌های مختلفی برای طبقه‌بندی داده ارائه شده است اما در این میان شبکه‌های عصبی مخاطبان بیشتری را به خود جذب کرده‌اند. مهم‌ترین مسئله در این نوع از طبقه‌بندی‌کننده‌ها انتخاب روشی مناسب برای آموزش شبکه‌های عصبی است. متداول‌ترین روش آموزشی شبکه‌های عصبی روش پس‌انتشار خطا است که همگرایی کند و توقف در نقاط بهینه محلی از مهم‌ترین نقاط ضعف آن محسوب می‌شود. رویکرد جدید محققین استفاده از الگوریتم‌های ابتکاری در فرایند آموزش شبکه‌های عصبی است. در این مقاله پیشنهاد استفاده از روش آموزشی نوینی به نام الگوریتم جستجوی گرانشی (GSA) در آموزش شبکه‌های عصبی به‌منظور طبقه‌بندی داده‌ها ارائه می‌شود. روش GSA آخرین و جدیدترین نسخه از انواع روش‌های بهینه‌سازی هوش جمعی است که با الهام از مفاهیم جرم و نیروی جاذبه و با شبیه‌سازی قوانین مرتبط با آن ارائه شده است. در این مقاله با استفاده از روش GSA یک شبکه عصبی MLP جهت طبقه‌بندی پنج مجموعه داده مرجع آموزش داده می‌شود. همچنین کارایی روش پیشنهادی در آموزش و آزمایش شبکه عصبی با دو روش آموزشی پس‌انتشار خطا و بهینه‌سازی گروه ذرات مقایسه می‌شود. نتایج نهایی نشان می‌دهند در اکثر موارد روش GSA قابلیت چشمگیری در طبقه‌بندی صحیح داده‌ها دارد. به‌علاوه در آزمایشات انجام‌گرفته ویژگی منحصر به‌فردی از روش GSA پدیدار شد و آن پایداری نسبتاً عالی در طبقه‌بندی صحیح داده‌ها در تمام موارد بود. از نقطه نظر معیار زمانی نیز روش GSA نسبت به روش PSO در زمان کمتری به پاسخ مناسب دست می‌یابد.

کلید واژه: الگوریتم‌های ابتکاری، الگوریتم جستجوی گرانشی، شبکه عصبی MLP، طبقه‌بندی داده.

۱- مقدمه

طبقه‌بندی^۱ یکی از حوزه‌های تحقیقاتی مهم در داده‌کاوی^۲ است. "طبقه‌بندی داده" یعنی اختصاص دادن یک بردار ویژگی ورودی به یکی از مجموعه‌های هدف [۱]. در این راستا روش‌های طبقه‌بندی متنوعی مانند درخت تصمیم [۲] و [۳]، شبکه‌های عصبی مصنوعی^۳ (ANN) [۴]، ماشین بردار پشتیبان [۵] و [۶]، نزدیک‌ترین همسایه، روش‌های هیستوگرام و هسته ارائه شده‌اند. در میان روش‌های موجود، شبکه‌های عصبی مصنوعی متداول‌ترین شیوه برای طبقه‌بندی داده هستند.

این مقاله در تاریخ ۱۱ مرداد ماه ۱۳۸۹ دریافت و در تاریخ ۱ خرداد ماه ۱۳۹۰ بازنگری شد.

مریم ده‌باشیان، گروه مخابرات و الکترونیک دانشکده مهندسی، دانشگاه بیرجند، (email: m.dehbashian@gmail.com)

سیدحمید ظهیری، استادیار گروه مخابرات و الکترونیک دانشکده مهندسی، دانشگاه بیرجند، (email: hzahiri@birjand.ac.ir)

1. Classification
2. Data Mining
3. Artificial Neural Network

4. Multi Layer Perceptron
5. Error Back Propagation

هوش جمعی است که از قوانین جاذبه میان اجرام و حرکت نیوتنی الهام گرفته شده است. طبق قانون جاذبه نیوتن، هر جسم به اجسام دیگر نیرو وارد نموده و آنها را به سمت خود جذب می‌کند. به وضوح هر چه این اجسام بزرگتر و نزدیک‌تر باشند، تأثیر این نیرو بیشتر خواهد بود. در نتیجه هر جسم با استفاده از نیروی جاذبه محل و مقدار جرم سایر اجسام را درک می‌کند، بنابراین می‌توان از این نیرو به‌عنوان رسانه‌ای برای تبادل اطلاعات استفاده کرد.

از الگوریتم جستجوی گرانشی در حل مسائل بهینه‌سازی استفاده می‌شود. در این الگوریتم پاسخ‌های مورد نظر موقعیت اجرام در فضای مسئله هستند، میزان اجرام نیز با توجه به تابع هدف تعیین می‌شود.

در ابتدا فضای سیستم مشخص می‌شود که شامل یک دستگاه مختصات چندبُعدی در فضای تعریف مسئله است. پس از تشکیل سیستم، قوانین حاکم بر آن مشخص می‌شوند. فرض می‌شود تنها قانون جاذبه و قوانین حرکت بر این سیستم حاکم هستند. صورت کلی این قوانین تقریباً شبیه قوانین طبیعت است و به‌صورت زیر تعریف می‌شوند:

سیستم به‌صورت مجموعه‌ای از m جرم تصور می‌شود. موقعیت هر جرم می‌تواند جوابی برای مسئله می‌باشد. موقعیت بُعد d از جرم i با x_i^d نشان داده می‌شود

$$X_i = (x_i^1, \dots, x_i^d, \dots, x_i^n) \quad \text{for } i = 1, 2, \dots, m \quad (1)$$

در رابطه فوق n نشان‌دهنده بُعد فضای پاسخ است.

در این سیستم در زمان t ، به هر جرم i از سوی جرم j جهت بُعد d ، نیرویی به اندازه $F_{ij}^d(t)$ وارد می‌شود. مقدار این نیرو طبق (۲) محاسبه می‌شود. $G(t)$ ثابت گرانش در زمان t و R_{ij} فاصله بین دو جرم i و j می‌باشد. برای تعیین فاصله بین اجرام مطابق (۳) از فاصله اقلیدسی (نرم ۲) استفاده می‌شود

$$F_{ij}^d(t) = G(t) \frac{M_i(t) \times M_j(t)}{R_{ij}(t) + \epsilon} (x_j^d(t) - x_i^d(t)) \quad (2)$$

$$R_{ij}(t) = \|X_i(t) - X_j(t)\| \quad (3)$$

در (۲)، ϵ یک عدد بسیار کوچک است. نیروی وارد بر جرم i در جهت بُعد d در زمان t $(F_i^d(t))$ ، برابر مجموع نیروهایی است که k جرم برتر جمعیت بر جرم وارد می‌کنند. مقصود از اجرام برتر، عامل‌هایی هستند که دارای برزندگی بیشتری باشند

$$F_i^d(t) = \sum_{j \in kbest, j \neq i} rand_j(t) \times F_{ij}^d(t) \quad (4)$$

در رابطه فوق $kbest$ بیانگر مجموعه k جرم برتر جمعیت است. همچنین $rand_j$ عددی تصادفی با توزیع یکنواخت در بازه $[0, 1]$ است که برای حفظ خصوصیت تصادفی بودن جستجو در نظر گرفته می‌شود.

طبق قانون دوم نیوتن، هر جرم در جهت بُعد d شتابی می‌گیرد که متناسب است با نیروی وارد بر جرم i در آن جهت بخش بر جرم i . رابطه (۵) شتاب جرم i در جهت بُعد d در زمان t را با $a_i^d(t)$ نشان می‌دهد

$$a_i^d(t) = \frac{F_i^d(t)}{M_i(t)} \quad (5)$$

سرعت هر جرم برابر مجموع ضربی از سرعت فعلی جرم و شتاب جرم طبق (۶) تعریف می‌شود. موقعیت جدید بُعد d از جرم i طی (۷) محاسبه می‌شود

$$V_i^d(t+1) = rand_i \times V_i^d(t) + a_i^d(t) \quad (6)$$

آنها به‌صورت جمعیتی عمل می‌کنند. روش‌های جستجوی ابتکاری بر خلاف روش‌های کلاسیک بر مبنای تصادف عمل کرده و جستجویی فضا را به‌صورت موازی انجام می‌دهند. این نوع روش‌ها تنها از تابع برزندگی برای هدایت جستجو استفاده می‌کنند.

از نمونه این الگوریتم‌ها، الگوریتم ژنتیک [۱۷] با الهام از علم وراثت و تکامل است که در [۱۸] برای محاسبه ضرایب شبکه عصبی از این الگوریتم استفاده شده است. در [۱۹] تا [۲۲] کارایی بهتر این روش در مقایسه با روش پسانتشار خطا و روش‌های مبتنی بر گرادیان نشان داده شده است. در [۲۳] با استفاده از این الگوریتم شبکه‌های عصبی MLP جهت طبقه‌بندی داده آموزش داده شده‌اند. الگوریتم ابتکاری دیگر، بازپخت شبیه‌سازی شده [۲۴] با الهام از مشاهدات ترمودینامیک است. در [۲۵] و [۲۶] کارایی بهتر این روش در مقایسه با روش پسانتشار خطا نشان داده شده است. در [۲۷] نیز آموزش شبکه عصبی MLP جهت طبقه‌بندی داده با الگوریتم بازپخت شبیه‌سازی شده گزارش شده است. الگوریتم سیستم ایمنی مصنوعی [۲۸] از دیگر الگوریتم‌های ابتکاری است که از سیستم ایمنی بدن انسان الهام گرفته شده است. در [۲۹] از این الگوریتم برای آموزش شبکه‌های عصبی MLP استفاده شده است. از الگوریتم‌های ابتکاری دیگر می‌توان به بهینه‌سازی گروه ذرات [۳۰] با تقلید از رفتار اجتماعی پرندگان و ماهی‌ها اشاره کرد. در [۳۱] تا [۳۴] کارایی برتر این روش در مقایسه با روش پسانتشار خطا نشان داده شده است. از این الگوریتم در [۳۵] و [۳۶] برای آموزش شبکه‌های عصبی MLP جهت طبقه‌بندی داده استفاده شده است. در [۳۷] نیز از این الگوریتم و الگوریتم جمعیت مورچگان برای آموزش شبکه‌های عصبی برای طبقه‌بندی استفاده شده است.

به تازگی در حوزه الگوریتم‌های هوش جمعی، روش بهینه‌سازی و جستجوی جدیدی به نام روش جستجوی گرانشی (GSA) معرفی شده است [۳۸]. این الگوریتم با الهام از مفاهیم جرم و نیروی جاذبه و با شبیه‌سازی قوانین مرتبط با آن ارائه شده است. این الگوریتم تا به حال در کاربردهای متنوعی همچون طراحی فیلتر [۳۹]، انتخاب ویژگی [۴۰]، جبران‌سازی توان راکتیو [۴۱] و ... استفاده شده است. در [۴۲] یک شبکه عصبی MLP با استفاده از روش GSA جهت فشرده‌سازی تصاویر آموزش داده شده است.

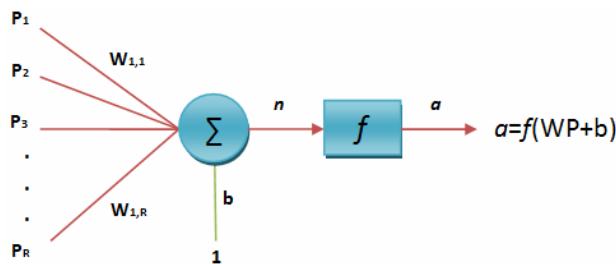
در این مقاله یک شبکه عصبی MLP جهت طبقه‌بندی داده‌ها با استفاده از الگوریتم جستجوی گرانشی آموزش داده شده و سپس از آن برای طبقه‌بندی پنج مجموعه داده مرجع استفاده می‌شود. در ادامه کارایی روش پیشنهادی در آموزش و آزمایش شبکه عصبی با دو روش آموزشی دیگر به نام‌های پسانتشار خطا و بهینه‌سازی گروه ذرات در طبقه‌بندی داده‌های پنج مجموعه مرجع مورد مقایسه قرار می‌گیرد.

در بخش ۲ مقاله الگوریتم جستجوی گرانشی معرفی می‌شود. در بخش ۳ نحوه آموزش شبکه عصبی MLP با استفاده از روش GSA معرفی خواهد شد. در بخش ۴ عملکرد طبقه‌بندی کننده طراحی شده ارزیابی می‌شود و بخش ۵ ضمن نتیجه‌گیری و بحث پیرامون نتایج به‌دست آمده مقاله را به انتها می‌رساند.

۲- معرفی الگوریتم جستجوی گرانشی

الگوریتم جستجوی گرانشی جدیدترین عضو خانواده الگوریتم‌های

1. Particle Swarm Optimization
2. Gravitational Search Algorithm



شکل ۲: یک پرسپترون با ورودی بایاس.

$$best(t) = \min_{j \in \{1, \dots, m\}} fit_j(t) \quad (11)$$

$$worst(t) = \max_{j \in \{1, \dots, m\}} fit_j(t) \quad (12)$$

در ابتدای تشکیل سیستم، هر جسم به صورت تصادفی در یک نقطه از فضا قرار می‌گیرد که جوابی از مسئله است. در هر لحظه از زمان، اجرام ارزیابی شده، سپس تغییر مکان هر جرم پس از محاسبه (۱) تا (۱۲) محاسبه می‌شود. پارامترهای سیستم نیز در هر مرحله به‌روز رسانی می‌شوند. شرط توقف می‌تواند پس از طی مدت زمان مشخصی تعیین شود [۳۸]. در شکل ۱ فلوجارت الگوریتم GSA نشان داده شده است.

ویژگی‌های مثبت الگوریتم GSA همچون همگرایی سریع، عدم توقف در بهینه‌های محلی، کاهش حجم محاسباتی نسبت به الگوریتم‌های تکاملی و عدم نیاز به حافظه در مقایسه با دیگر الگوریتم‌های خانواده هوش جمعی، بستر جدیدی از تحقیقات را فراروی محققین قرار داده است. از این رو با توجه به زمینه‌های کاربردی مورد استفاده، نسخه‌های متفاوتی از این الگوریتم ارائه شده است که می‌توان به الگوریتم جستجوی گرانشی باینری^۳ (BGS) [۴۳]، الگوریتم جستجوی گرانشی نخبه‌گرای پیشرفته^۴ (AEGSA) [۴۴] و الگوریتم جستجوی گرانشی چندهدفه^۵ (MOGSA) [۴۵] اشاره نمود.

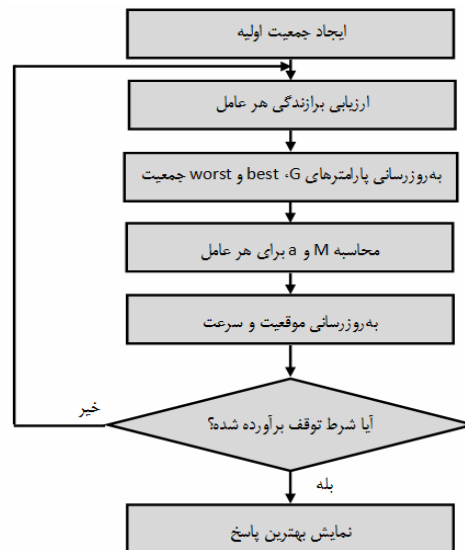
۳- آموزش شبکه عصبی MLP با استفاده از روش GSA

۳-۱ معرفی شبکه عصبی MLP

مفهوم پرسپترون در سال ۱۹۴۳ توسط McCulloch و Pitts به‌عنوان یک نرون مصنوعی ارائه شد.

یک شبکه پرسپترون چندلایه (MLP)، ارتباط غیر خطی میان بردارهای ورودی و خروجی را نشان می‌دهد. این عمل از طریق اتصال نرون‌های هر گره به لایه‌های قبلی و بعدی انجام می‌شود. خروجی هر نرون در ضرایب وزنی ضرب و به تابع تحریک غیر خطی به‌عنوان ورودی داده می‌شود. در مرحله آموزش به پرسپترون اطلاعات آموزشی داده می‌شود، سپس وزن‌های شبکه به‌گونه‌ای تنظیم می‌شوند که خطای بین خروجی فعلی و هدف حداقل گردند و یا این که تعداد دفعات آموزش، به مقداری که از پیش تعیین شده است، برسد. پس از آن برای ارزیابی صحت فرایند آموزش، یک سری ورودی‌های تجربه‌نشده به شبکه اعمال می‌شود. این ورودی‌ها متفاوت از ورودی‌های به‌کار رفته در فرایند آموزش شبکه هستند. عموماً آموزش شبکه‌های عصبی بسیار پیچیده و یک مسئله بهینه‌سازی با تعداد متغیر زیاد است [۴۶].

در شکل ۲ یک نرون با چندین ورودی نشان داده شده است. در اینجا



شکل ۱: فلوجارت الگوریتم جستجوی گرانشی.

$$x_i^d(t+1) = x_i^d(t) + V_i^d(t+1) \quad (7)$$

در روابط فوق $V_i^d(t)$ سرعت بُعد d عامل i در زمان t و $rand_i$ عددی تصادفی با توزیع یکنواخت در بازه $[0,1]$ است که برای حفظ خصوصیت تصادفی بودن جستجو در نظر گرفته می‌شود. برای تنظیم ضریب گرانش از (۸) استفاده می‌شود

$$G(t) = \beta^{-\frac{t}{T}} \quad (8)$$

ثابت گرانش پارامتری مناسب برای کنترل دو ویژگی کاوش^۱ و بهره‌وری^۲ در این الگوریتم به شمار می‌آید. مقادیر بزرگ آن موجب تقویت توانایی کاوش الگوریتم و مقادیر کوچک آن موجب افزایش توانایی بهره‌وری الگوریتم می‌شود. از آنجا که در مراحل اولیه جستجو لازم است الگوریتم به جستجوی نقاط جدیدی در فضای مسئله پرداخته و در مراحل پایانی با افزایش توان بهره‌وری به بهبود جواب‌های دیده‌شده بپردازد، گزینه مناسب برای ثابت گرانش بایستی با یک مقدار اولیه بزرگ شروع شده و با گذشت زمان مقدار آن به تدریج کاهش یابد. طبق آزمایش‌های متعدد انجام‌شده، استفاده از رابطه‌نمایی جهت کاهش ثابت گرانش در حل بسیاری از مسائل مؤثر است.

در (۹) جرم عامل‌ها بر مبنای تابع هدف آنها تنظیم می‌شود به‌گونه‌ای که به عامل‌های با شایستگی بیشتر، جرم بیشتری نسبت داده می‌شود

$$m_i(t) = \frac{fit_i(t) - worst(t)}{best(t) - worst(t)} \quad (9)$$

در این رابطه $fit_i(t)$ بیانگر میزان برازندگی جرم i در زمان t است، $best(t)$ و $worst(t)$ به‌ترتیب بیانگر میزان شایستگی قوی‌ترین و ضعیف‌ترین عامل جمعیت در زمان t هستند. در نهایت اندازه جرم عامل‌ها طبق (۱۰) نرمالیزه می‌شود

$$M_i(t) = \frac{m_i(t)}{\sum_{j=1}^N m_j(t)} \quad (10)$$

در مسائل کمینه‌یابی می‌توان از روابط زیر برای محاسبه بهترین و بدترین عامل‌ها استفاده کرد

3. Binary Gravitational Search Algorithm

4. Advanced Elitism Gravitational Search Algorithm

5. Multi Objective Gravitational Search Algorithm

1. Exploration

2. Exploitation

به روز رسانی وزن‌های شبکه در جهت کاهش MSE در تکرار جاری است. در هر تکرار، تمام اجرام موقعی‌شان را با محاسبه سرعت جدید به روز رسانی و به سمت موقعیت جدید حرکت می‌کنند. موقعیت جدید مجموعه‌ای از وزن‌های جدید است که برای به دست آوردن MSE جدید استفاده می‌شود. این فرایند برای تمام اجرام تکرار می‌شود. جرم با حداقل MSE به عنوان بهترین جرمی که تاکنون پیدا شده، در نظر گرفته می‌شود. فرایند آموزش تا زمانی ادامه پیدا می‌کند که خطای قابل قبول توسط یک جرم به دست آید یا یک شرط خاتمه مانند تعدادی تکرار پیش فرض، برآورده شود. زمانی که فرایند آموزش خاتمه یافت، به ازای شایسته‌ترین جرم (حداقل MSE)، بردار وزن متناظر با آن استخراج می‌شود.

۴- ارزیابی عملکرد

برای ارزیابی عملکرد شبکه عصبی MLP در طبقه‌بندی داده‌ها، از آن برای طبقه‌بندی پنج مجموعه داده مرجع استفاده می‌شود. پنج مجموعه داده از وب سایت مجموعه داده‌های یادگیری ماشین دانشگاه کالیفرنیا [۴۷] انتخاب شده است:

۱) داده‌های انواع گل زنبق^۱

داده‌های گل زنبق شامل ۵۰ اندازه‌گیری از چهار ویژگی سه نوع گل زنبق به نام‌های Setosa، Versicolor و Vergenica است. این ویژگی‌ها عبارتند از طول کاسبرگ، عرض کاسبرگ، طول گلبرگ و عرض گلبرگ. بنابراین در این مسئله سه کلاس موجود است و هر کلاس دارای ۵۰ الگو با بردارهای ویژگی چهاربندی است.

۲) داده‌های انواع شراب^۲

این مجموعه متشکل از آنالیز شیمیایی عناصر رشد داده شده شراب‌ها در شرایط مساوی در منطقه‌ای در ایتالیا است. بردارهای ویژگی دارای ۱۳ مؤلفه پیوسته هستند. در مجموع ۱۷۸ نمونه در سه کلاس مجزا در این مجموعه داده وجود دارد.

۳) داده‌های سرطان سینه^۳

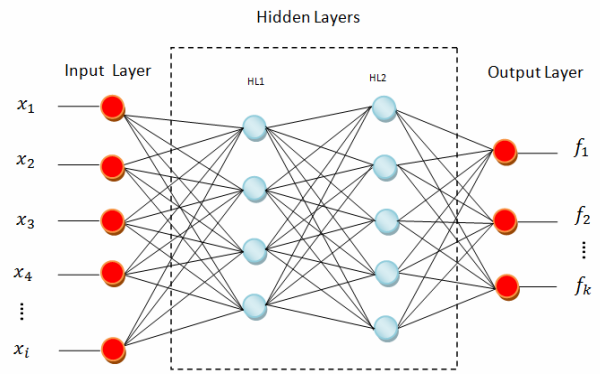
این مجموعه شامل ۶۹۹ نمونه مربوط به توده‌های سرطانی موجود در سینه بیماران دانشگاه Wisconsin است. نمونه‌ها دارای ابعاد ۹بندی بوده و در دو کلاس خوش خیم و بدخیم طبقه‌بندی شده‌اند.

۴) داده‌های ارزیابی دستیاران آموزشی^۴ (TAE)

این داده‌ها شامل ارزیابی عملکرد تدریس بیش از سه ترم پی‌درپی و دو ترم تابستانی ۱۵۱ دستیار آموزشی در دانشگاه Wisconsin - Madison در گروه آمار است. هر نمونه دارای ابعاد ۵بندی بوده و در سه کلاس ضعیف، متوسط و عالی برای ارزیابی تدریس هر دستیار آموزشی طبقه‌بندی شده‌اند.

۵) داده‌های تشخیص بیماری قلبی از روی تصاویر SPECT^۵

این مجموعه شامل ۲۶۷ تصویر SPECT از بیماران قلبی است. بردارهای ویژگی دارای ۲۲ مؤلفه باینری هستند و بیماران در دو کلاس نرمال و غیر نرمال طبقه‌بندی می‌شود. ۷۰٪ از نقاط مجموعه داده ۱ تا ۴ به صورت کاملاً تصادفی انتخاب و به عنوان نقاط آموزشی برای شبکه عصبی در نظر گرفته می‌شوند. از ۳۰٪



شکل ۳: نمایش ساختار شبکه عصبی MLP.

ورودی‌های P_1, P_2, \dots, P_R در ماتریس W که همان ماتریس وزن‌ها می‌باشد، ضرب شده و سپس در جمع‌کننده (Σ) جمع می‌شوند، حاصل جمع همراه با نرون بایاس b وارد تابع تحریک f می‌شود

$$n = W_{1,1}P_1 + W_{1,2}P_2 + \dots + W_{1,R}P_R + b \quad (13)$$

رابطه فوق را می‌توان به صورت ساده‌تر نیز نوشت

$$n = W \times P + b \quad (14)$$

۳-۲ طراحی شبکه عصبی MLP

در این تحقیق برای طبقه‌بندی داده‌ها از شبکه عصبی MLP استفاده می‌شود. شبکه عصبی MLP از لایه ورودی، لایه‌های مخفی و لایه خروجی که هر یک شامل تعداد نرون‌های مشخصی است تشکیل می‌شود. تعداد نرون‌های لایه ورودی برابر تعداد ویژگی‌های بردار ورودی به علاوه یک نرون بایاس، و تعداد نرون‌های لایه خروجی برابر تعداد کلاس‌های تعریف شده برای طبقه‌بندی کننده است.

انتخاب تعداد لایه‌های مخفی و تعداد نرون‌ها در هر یک از آنها بسیار مهم است، چرا که اگر تعداد آنها کم باشد شبکه برای حل مسائل غیر خطی و پیچیده با کمبود منابع یادگیری مواجه می‌شود و اگر زیاد باشد موجب ایجاد دو مشکل خواهد شد. اول آن که زمان آموزش شبکه افزایش می‌یابد و دوم آن که ممکن است شبکه نظام بی‌اهمیت داده‌های آموزش را یاد بگیرد و در حل مسائل ضعیف عمل نماید.

از این رو با توجه به بهترین نتایج تجربی به دست آمده برای شبکه مورد نظر دو لایه مخفی در نظر گرفته شد، لایه مخفی اول دارای ۴ نرون و لایه مخفی دوم دارای ۵ نرون است. ساختار شبکه عصبی MLP در شکل ۳ نشان داده شده است.

۳-۳ آموزش شبکه عصبی MLP توسط روش GSA

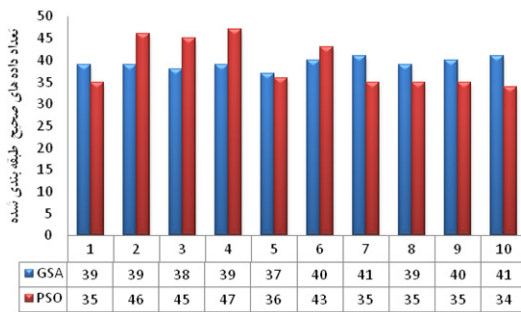
در فرایند آموزش شبکه عصبی MLP جهت طبقه‌بندی داده‌ها، هدف نهایی جستجوی بهترین اوزان شبکه عصبی است که به کوچک‌ترین میانگین مربعات خطا (MSE) منتهی شود، در این تحقیق روش GSA برای یافتن بهترین اوزان شبکه عصبی پیشنهاد می‌شود.

با توجه به تعداد کل نرون‌های شبکه عصبی MLP (شکل ۳)، بردار وزنی به شکل تصادفی تولید می‌شود

$$W = [w_1, w_2, w_3, \dots, w_n] \quad (15)$$

در رابطه فوق n تعداد نرون‌های موجود در شبکه عصبی MLP است. در روش GSA جمعیت اولیه‌ای از اجرام (در اینجا بردارهای وزن همان اجرام هستند) تولید می‌شود. اجرام در فضا حرکت می‌کنند و سعی می‌کنند خطا را مینیمم سازند. تغییر موقعیت یک جرم به معنای

1. Iris
2. Wine
3. Breast Cancer
4. Teaching Assistant Evaluation
5. Single Proton Emission Computed Tomography



شکل ۵: مقایسه تعداد داده‌های صحیح در طبقه‌بندی شراب توسط دو روش GSA و PSO در مرحله آزمایش شبکه عصبی.

جدول ۲: بررسی نرخ تشخیص صحیح داده‌های آموزشی و آزمایشی شبکه عصبی MLP در طبقه‌بندی داده‌های شراب توسط سه روش آموزشی.

Wine	درصد طبقه‌بندی صحیح	GSA	PSO	EBP
Train Stage	Min	%۷۴٫۴۰	%۶۸٫۸۰	%۴۰٫۸۰
	Max	%۷۶٫۰۰	%۸۶٫۴۰	%۴۵٫۶۰
Test Stage	Average	۷۵٫۶۰	%۷۸٫۱۶	%۴۲٫۷۲
	Variance	%۰٫۰۰۳۲۰	%۰٫۵۰۰۶۹	%۰٫۰۱۷۳۵
	Time (Sec)	۲۱۰	۲۹۰٫۴	۷٫۶
		Min	%۶۹٫۸۱	%۶۴٫۱۵
Test Stage	Max	%۷۷٫۳۶	%۸۸٫۶۸	%۴۱٫۵۱
	Average	%۷۴٫۱۵	%۷۳٫۷۷	%۳۳٫۵۸
	Variance	%۰٫۰۵۵۷۷	%۱٫۰۳۹۹۱	%۰٫۱۴۸۷۳

کتر باشد نشان‌دهنده این است که روش مورد استفاده از پراکندگی و اختلاف کمتری در تشخیص صحیح برخوردار است. با توجه به نتایج جدول ۱، روش GSA نسبت به دو روش دیگر عملکرد بهتری در تشخیص صحیح داده‌های آموزشی و آزمایشی شبکه عصبی MLP در طبقه‌بندی داده‌های گل زنبق دارد.

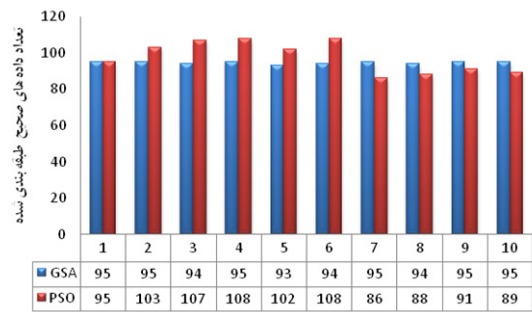
۴-۲ ارزیابی عملکرد در طبقه‌بندی داده‌های Wine

در جدول ۲ نرخ تشخیص صحیح داده‌های آموزشی و آزمایشی شبکه عصبی MLP توسط سه روش آموزشی در طبقه‌بندی داده‌های شراب بر حسب درصد مورد بررسی قرار گرفته است.

با توجه به نتایج جدول ۲، روش GSA نسبت به روش متداول پس‌انتشار خطا عملکرد بسیار بهتری دارد. از سوی دیگر روش PSO نسبت به GSA در طبقه‌بندی این داده‌ها دارای نرخ تشخیص صحیح بالاتری است. اما قابل توجه است که در روش GSA بر خلاف دو روش دیگر میانگین نرخ تشخیص صحیح با مقادیر کمینه و بیشینه آن فاصله کمی دارد، این موضوع با توجه به مقدار کم واریانس نیز قابل استنباط است. به بیان روشن‌تر می‌توان گفت الگوریتم GSA ثبات و پایداری خوبی در تشخیص صحیح داده‌ها دارد، از این رو این خصوصیت مهم روش GSA را "امنیت پاسخ" می‌نامیم. شکل‌های ۴ و ۵ مقایسه‌ای را میان دو روش GSA و PSO از نقطه نظر تعداد داده‌های صحیح طبقه‌بندی شده در مراحل آموزش و آزمایش نشان می‌دهند (تعداد کل داده‌ها در مرحله آموزش ۱۲۵ عدد و در مرحله آزمایش ۵۳ عدد است).

۴-۳ ارزیابی عملکرد در طبقه‌بندی داده‌های Breast Cancer

در جدول ۳ نرخ تشخیص صحیح داده‌های آموزشی و آزمایشی شبکه عصبی MLP توسط سه روش آموزشی در طبقه‌بندی داده‌های سرطان



شکل ۴: مقایسه تعداد داده‌های صحیح در طبقه‌بندی داده‌های شراب توسط دو روش GSA و PSO در مرحله آموزش شبکه عصبی.

جدول ۱: بررسی نرخ تشخیص صحیح داده‌های آموزشی و آزمایشی شبکه عصبی MLP در طبقه‌بندی داده‌های گل زنبق توسط سه روش آموزشی.

Iris	درصد طبقه‌بندی صحیح	GSA	PSO	EBP
Train Stage	Min	%۹۷٫۱۴	%۶۳٫۸۱	%۳۵٫۲۴
	Max	%۹۸٫۱۰	%۱۰۰٫۰۰	%۹۹٫۰۵
Test Stage	Average	%۹۸٫۰۰	%۹۴٫۸۶	%۸۷٫۸۱
	Variance	%۰٫۰۰۰۹	%۱٫۲۰۱۷	%۳٫۴۸۰۶
	Time (Sec)	۱۳۷٫۷	۴۷۵٫۷	۶٫۵
		Min	%۹۲٫۳۳	%۶۴٫۴۴
Test Stage	Max	%۱۰۰٫۰۰	%۹۷٫۷۸	%۹۵٫۵۶
	Average	%۹۶٫۲۲	%۹۱٫۳۳	%۸۷٫۵۶
	Variance	%۰٫۰۷۷۴	%۰٫۹۳۷۷	%۴٫۲۶۹۸

باقیمانده به‌عنوان نقاط آزمایشی استفاده خواهد شد. همچنین مطابق با وب سایت مجموعه داده‌های یادگیری ماشین، ۳۰٪ از نقاط مجموعه داده ۵ به آموزش و ۷۰٪ باقیمانده به آزمایش اختصاص داده می‌شود. در شرایط کاملاً یکسان شبکه عصبی مورد نظر جهت طبقه‌بندی داده‌ها توسط روش‌های GSA، PSO و EBP آموزش داده می‌شود. تمام نتایج برای ۱۰۰۰ بار تکرار هر الگوریتم ثبت می‌شود، همچنین با توجه به تصادفی بودن روش‌های فوق، تمام آزمایشات برای ۱۰ مرتبه تکرار انجام می‌شود. مشخصات روش‌های به کار رفته برای آموزش به شرح زیر است:

- روش GSA

جمعیت اولیه اجرام (بردارهای وزن) در روش GSA برابر ۲۰ می‌باشد. در (۸) مقدار α با توجه به نتایج تجربی برابر ۲۰ و مقدار β به صورت خطی از ۱ تا ۳ افزایش می‌یابد.

- روش PSO

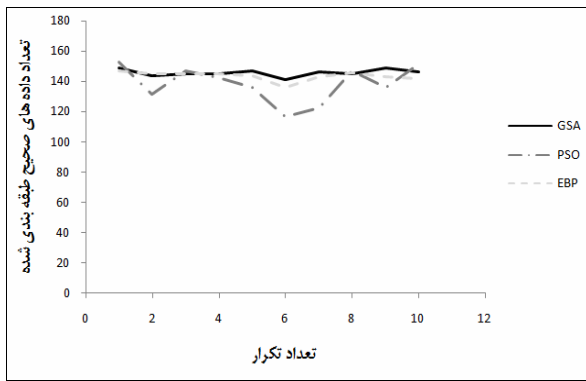
جمعیت اولیه ذرات (بردارهای وزن) برابر ۲۰ می‌باشد، در این روش فاکتورهای یادگیری C_1 و C_2 با هم برابر و مساوی ۲ فرض می‌شوند. همچنین وزن اینرسی W به صورت خطی از ۰٫۹ تا ۰٫۲ کاهش می‌یابد.

- روش پس‌انتشار خطا (EBP)

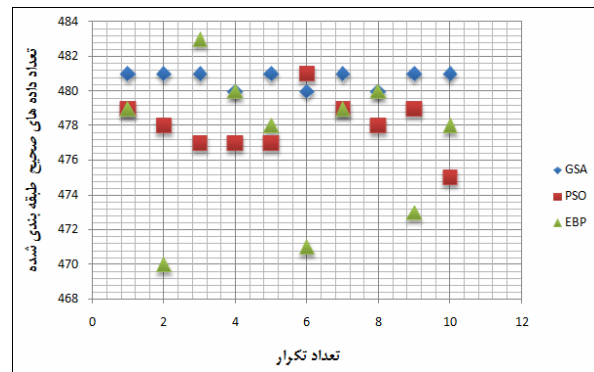
در این روش $\eta = ۰٫۸$ فرض شده است.

۴-۱ ارزیابی عملکرد در طبقه‌بندی داده‌های Iris

در جدول ۱ نرخ تشخیص صحیح داده‌های آموزشی و آزمایشی شبکه عصبی MLP توسط سه روش آموزشی در طبقه‌بندی داده‌های گل زنبق بر حسب درصد مورد بررسی قرار گرفته است. در این جدول سطر Min، Max، Average و Variance به ترتیب نشان‌دهنده کمترین، بیشترین، میانگین و واریانس نرخ تشخیص صحیح در ده مرتبه اجرای الگوریتم طبقه‌بندی داده‌های گل زنبق بر حسب درصد است. هرچه مقدار واریانس



شکل ۷: مقایسه پراکندگی تعداد داده‌های صحیح در طبقه‌بندی داده‌های SPECT توسط سه روش آموزشی در مرحله آزمایش شبکه عصبی.



شکل ۶: مقایسه پراکندگی تعداد داده‌های صحیح در طبقه‌بندی داده‌های سرطان سینه توسط سه روش آموزشی در مرحله آموزش شبکه عصبی.

جدول ۵: بررسی نرخ تشخیص صحیح داده‌های آموزشی و آزمایشی شبکه عصبی MLP در طبقه‌بندی داده‌های تصاویر SPECT توسط سه روش آموزشی.

SPECT	درصد طبقه‌بندی صحیح	GSA	PSO	EBP
Train Stage	Min	%۹۰٫۰۰	%۸۷٫۵۰	%۸۸٫۷۵
	Max	%۹۲٫۵۰	%۹۲٫۵۰	%۹۲٫۵۰
	Average	%۹۱٫۵۰	%۹۰٫۱۳	%۹۰٫۰۰
	Variance	%۰٫۰۰۶۲۵	%۰٫۰۳۹۶۷	%۰٫۰۱۷۳۶
	Time (Sec)	۱۷۹٫۶	۱۹۳	۹٫۱
Test Stage	Min	%۷۵٫۴۰	%۶۲٫۵۷	%۷۲٫۷۳
	Max	%۷۹٫۶۸	%۸۱٫۸۲	%۷۸٫۶۱
	Average	%۷۷٫۹۱	%۷۴٫۱۲	%۷۶٫۷۹
	Variance	%۰٫۰۱۵۹۲	%۰٫۴۱۷۶۴	%۰٫۰۲۶۸۲

جدول ۳: بررسی نرخ تشخیص صحیح داده‌های آموزشی و آزمایشی شبکه عصبی MLP در طبقه‌بندی داده‌های سرطان سینه توسط سه روش آموزشی.

Breast Cancer	درصد طبقه‌بندی صحیح	GSA	PSO	EBP
Train Stage	Min	%۹۸٫۱۶	%۹۷٫۱۴	%۹۶٫۱۱
	Max	%۹۸٫۳۶	%۹۸٫۳۶	%۹۸٫۷۷
	Average	%۹۸٫۳۰	%۹۷٫۷۵	%۹۷٫۵۷
	Variance	%۰٫۰۰۰۱۰	%۰٫۰۰۱۱۲	%۰٫۰۰۷۶۶
	Time (Sec)	۶۸۸	۶۹۲	۲۲٫۵
Test Stage	Min	%۹۵٫۷۱	%۹۴٫۷۶	%۹۳٫۸۱
	Max	%۹۷٫۶۲	%۹۷٫۶۲	%۹۶٫۱۹
	Average	%۹۶٫۴۳	%۹۶٫۲۴	%۹۶٫۸۶
	Variance	%۰٫۰۰۵۶۷	%۰٫۰۱۰۳۰	%۰٫۰۰۶۴۵

طبقه‌بندی داده‌ها ارائه می‌دهد.

۴- ارزیابی عملکرد در طبقه‌بندی داده‌های TAE

در جدول ۴ نرخ تشخیص صحیح داده‌های آموزشی و آزمایشی شبکه عصبی MLP توسط سه روش آموزشی در طبقه‌بندی داده‌های ارزیابی دستیاران آموزشی بر حسب درصد مورد بررسی قرار گرفته است. با توجه به نتایج جدول ۴، روش GSA نسبت به دو روش دیگر نتایج برتری در دو مرحله آموزش و آزمایش شبکه عصبی MLP در طبقه‌بندی این نوع داده‌ها دارد. همچنین با توجه به معیار واریانس، نتایج به دست آمده توسط روش GSA دارای امنیت پاسخ بالاتری می‌باشد.

۵- ارزیابی عملکرد در طبقه‌بندی داده‌های SPECT

در جدول ۵ نرخ تشخیص صحیح داده‌های آموزشی و آزمایشی شبکه عصبی MLP توسط سه روش آموزشی در طبقه‌بندی داده‌های تصاویر SPECT بر حسب درصد مورد بررسی قرار گرفته است. با توجه به نتایج جدول ۵، روش GSA رقابت نزدیکی با دو روش آموزشی دیگر دارد اما در مجموع و با توجه به معیار واریانس گزارش شده، الگوریتم GSA عملکرد بهتری در طبقه‌بندی داده‌های تصاویر SPECT در دو مرحله آموزش و آزمایش شبکه عصبی MLP دارد.

شکل ۷ مقایسه‌ای را میان تعداد داده‌های صحیح در طبقه‌بندی داده‌های SPECT توسط سه روش آموزشی برای ۱۰ مرتبه اجرای الگوریتم نشان می‌دهد. این مقایسه برای مرحله آزمایش شبکه عصبی انجام شده و تعداد داده‌های آزمایشی ۱۸۷ عدد است. با توجه به این شکل به روشنی مشخص است فاصله پراکندگی داده‌های طبقه‌بندی شده توسط روش GSA نسبت به دو روش دیگر بسیار کمتر است.

جدول ۴: بررسی نرخ تشخیص صحیح داده‌های آموزشی و آزمایشی شبکه عصبی MLP در طبقه‌بندی داده‌های TAE توسط سه روش آموزشی.

TAE	درصد طبقه‌بندی صحیح	GSA	PSO	EBP
Train Stage	Min	%۶۵٫۰۹	%۴۹٫۰۶	%۲۳٫۵۸
	Max	%۶۷٫۹۲	%۶۳٫۲۱	%۵۰٫۹۴
	Average	۶۶٫۴۲	%۵۶٫۸۹	%۴۰٫۷۵
	Variance	%۰٫۰۰۸۳۱	%۰٫۱۷۸۱۰	%۰٫۵۸۸۹۸
	Time (Sec)	۱۴۷	۱۸۱	۶٫۶
Test Stage	Min	%۵۱٫۱۱	%۳۵٫۵۶	%۲۰٫۰۰
	Max	%۶۰٫۰۰	%۴۶٫۶۷	%۴۰٫۰۰
	Average	%۵۷٫۵۶	%۴۱٫۷۸	%۳۲٫۴۴
	Variance	%۰٫۱۲۵۶۵	%۰٫۱۵۱۴۴	%۰٫۴۴۱۱۵

سینه بر حسب درصد مورد بررسی قرار گرفته است. در شکل ۶ نیز مقایسه‌ای میان تعداد داده‌های صحیح در طبقه‌بندی داده‌های سرطان سینه توسط سه روش آموزشی برای ۱۰ مرتبه اجرای الگوریتم نشان داده شده است. این مقایسه برای مرحله آموزش شبکه عصبی انجام شده و تعداد داده‌های آموزشی ۴۸۹ عدد است.

با توجه به نتایج جدول ۳، روش GSA نسبت به دو روش دیگر نتایج بهتری در دو مرحله آموزش و آزمایش شبکه عصبی MLP در طبقه‌بندی این نوع داده‌ها دارد. اما با توجه به نزدیکی نتایج به یکدیگر به خصوص در معیار میانگین نرخ تشخیص صحیح، شاید به اشتباه به نظر برسد تفاوت چندانی میان روش‌های آموزشی به کار رفته نیست. با توجه به معیار واریانس در جدول ۳ و فاصله پراکندگی داده‌های صحیح در شکل ۶ به وضوح مشخص است که مجموعه پاسخ‌های گزارش شده توسط روش GSA به یکدیگر نزدیک‌تر بوده و در مجموع امنیت پاسخ بالاتری را در

۴-۶ مقایسه زمان اجرای الگوریتم توسط سه روش آموزشی

در جداول ۱ تا ۵ زمان اجرای الگوریتم آموزش شبکه عصبی MLP در طبقه‌بندی پنج مجموعه داده مرجع توسط سه روش آموزشی به کار رفته آمده است. با توجه به نتایج ارائه‌شده، روش پس‌انتشار خطا به دلیل ساختار ساده و حجم محاسباتی بسیار کم دارای کمترین زمان اجرا است. در مقایسه زمان اجرای الگوریتم در دو روش آموزشی GSA و PSO، در تمامی موارد روش GSA به زمان کمتری نیاز دارد. با توجه به این که روش GSA دارای حجم محاسباتی بالاتری نسبت به روش PSO است، این کاهش زمان به علت بی‌حافظه بودن روش GSA است.

۵- نتیجه‌گیری

شبکه‌های عصبی مصنوعی متداول‌ترین شیوه برای طبقه‌بندی داده هستند. برای آموزش شبکه‌های عصبی تاکنون الگوریتم‌های آموزشی مختلفی ارائه شده‌اند. در این مقاله پیشنهاد استفاده از روش GSA برای آموزش شبکه عصبی به‌منظور طبقه‌بندی داده ارائه شده است. جهت بررسی صحت عملکرد روش پیشنهادی از آن برای آموزش یک شبکه عصبی MLP در طبقه‌بندی پنج مجموعه داده مرجع استفاده شده است. همچنین مقایسه‌ای میان روش آموزشی GSA و دو روش آموزشی دیگر به نام‌های PSO و EBP در طبقه‌بندی این مجموعه‌های داده صورت گرفته است. نتایج ارائه‌شده نشان دادند در اکثر موارد روش GSA در دو مرحله آموزش و آزمایش شبکه عصبی دارای قابلیت بهتری در طبقه‌بندی صحیح داده‌ها است. به‌علاوه ویژگی منحصر به فردی از روش GSA پدیدار شد و آن تولید پاسخ‌هایی با پایداری نسبتاً عالی بود. این ویژگی برای طبقه‌بندی‌کننده‌ها به‌ویژه زمانی که از آنها در کاربردهای حساس همچون پزشکی یا نظامی استفاده می‌شود، بسیار باارزش بوده و ضریب امنیت پاسخ سیستم طبقه‌بندی‌کننده را بالا می‌برد.

مراجع

- [1] K. Balakrishnan and V. Honavar, "Improving convergence of back propagation by handling flat - spots in the output layer," in *Proc. Second Int. Conf. on Artificial Neural Networks*, Brighton, UK, vol. 1, pp. 139-144, 1992.
- [2] M. T. Hagan and M. Menhaj, "Training feed forward networks with the Marquardt algorithm," *IEEE Trans. on Neural Networks*, vol. 5, no. 6, pp. 989-993, Nov. 1994.
- [3] W. Yan and S. Hongbao, "Improvement of neural network learning algorithm and its application in control," in *Proc. 3rd World Congress on Neural Networks*, vol. 2, pp. 971-975, Hefei, Anhui, Jul. 2000.
- [4] B. Bazartseren, G. Hildebrandt, and K. P. Holz, "Short - term water level prediction using neural networks and neuro - fuzzy approach," *Neuro Computing*, vol. 55, no. 3-4, pp. 439-450, Oct. 2003.
- [5] M. Engin, "ECG beat classification using neuro - fuzzy network," *Pattern Recognition Letters*, vol. 25, no. 15, pp. 1715-1722, Nov. 2004.
- [6] P. Kumar, S. N. Merchant, and U. B. Desai, "Improving performance in pulse radar detection using Bayesian regularization for neural network training," *Digital Signal Processing*, vol. 14, no. 5, pp. 438-448, Jul. 2004.
- [7] K. S. Tang, K. F. Man, and S. Kwong, Q. He, "Genetic algorithms and their applications," *IEEE Signal Processing Magazine*, vol. 13, no. 6, pp. 22-37, Nov. 1996.
- [8] L. L. Rogers, F. U. Dowla, and V. M. Johnson, "Optimal field - scale groundwater remediation using neural networks and the genetic algorithm," *Environmental Science and Technology*, vol. 29, no. 5, pp. 1145-1155, 1995.
- [9] D. J. Montana and L. Davis, "Training feed forward neural networks using genetic algorithms," *Machine Learning*, vol. 1, pp. 762-767, Aug. 1989.
- [10] J. Branke, "Genetic algorithms for neural network design and training," in *Proc. of the First Nordic Workshop on Genetic Algorithms*, vol. 1, pp. 145-163, University of Vaasa, Finland, Jan. 2003.
- [11] A. Ullah Khan, T. K. Bandopadhyaya, and S. Sharma, "Genetic algorithm based back propagation neural network performs better than back propagation neural network in stock rates prediction," *Inter. J. of Computer Science and Network Security*, vol. 8, no. 7, pp. 162-166, Jul. 2008.
- [12] L. A. Tarca, B. P. A. Grandjean, and F. Larachi, "Embedding monotonicity and concavity in the training of neural networks by means of genetic algorithms application to multiphase flow," *Computers and Chemical Engineering*, vol. 28, no. 9, pp. 1701-1713, 2004.
- [13] R. S. Sexton and R. E. Dorsey, "Reliable classification using neural network: a genetic algorithm and back propagation computation," *Decision Support Systems*, vol. 30, no. 1, pp. 11-22, Dec. 2000.
- [14] S. Kirkpatrick, C. D. Gelatto, and M. P. Vecchi, "Optimization by simulated annealing," *Science, New Series*, vol. 220, no. 4598, pp. 671-680, May. 1983.
- [15] A. L. Arnaud, P. J. L. Adeodato, G. C. Vasconcelos, and R. F. O. Neto, "MLP neural networks optimization through simulated annealing in a hybrid approach for time series prediction," *Congresso de Sociedade Brasileira de Computação*, pp. 1110-1113, Jul. 2005.
- [16] R. S. Sexton, R. E. Dorsey, and J. D. Johnson, "Optimization of neural networks: a comparative analysis of the genetic algorithm and simulated annealing," *European J. of Operational Research*, vol. 114, no. 3, pp. 589-601, May 1999.
- [17] S. W. Lin, T. Y. Tseng, S. Y. Chou, and S. C. Chen, "A simulated-annealing - based approach for simultaneous parameter optimization and feature selection of back - propagation networks," *Expert Systems with Applications*, vol. 34, no. 2, pp. 1491-1499, Feb. 2008.
- [18] J. D. Farmer, N. Packard, and A. Perelson, "The immune system, adaptation and machine learning," *Physica D*, vol. 2, no. 1-3, pp. 187-204, 1986.
- [19] R. Panti and L. N. De Castro, "The influence of diversity in an immune-based algorithm to train MLP networks," in *Proc. of the 6th Inter. Conf. on Artificial Immune Systems, ICARIS'07*, vol. 4628, pp. 71-82, 2007.
- [20] J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in *Proc. of IEEE Int. Conf. on Neural Networks*, vol. 4, pp. 1942-1948, 1995.
- [21] C. Marcio and B. L. Teresa, "An analysis of PSO hybrid algorithms for feed-forward neural networks training," in *Proc. of the Ninth Brazilian Symp. on Neural Networks, SBRN'06*, pp. 2-7, 2006.
- [22] F. A. Guerra and L. D. S. Coelho, "Multi - step ahead nonlinear identification of Lorenz's chaotic system using radial basis neural network with learning by clustering and particle swarm
- [1] R. S. Kulkarni and M. Vidyasagar, "Learning decision rules for pattern classification under a family of probability measures," *IEEE Trans. on Information Theory*, vol. 43, no. 1, pp. 154-166, Jan. 1997.
- [2] J. R. Qianlan, "Introduction of decision trees," *Machine Learning*, vol. 1, no. 1, pp. 86-101, Mar. 1986.
- [3] Y. Freund, "Boosting a weak learning algorithm by majority," *Information Computation*, vol. 121, no. 2, pp. 256-285, Sep. 1995.
- [4] C. Hongjun, R. Setiono, and L. Huan, "Effect data mining using neural networks," *IEEE Trans. on Knowledge and Data Engineering*, vol. 8, no. 6, pp. 957-961, Dec. 1996.
- [5] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proc. 5th Annual ACM Workshop on Computational Learning Theory*, New York, NY, US, pp. 144-152, 13-15 Jun. 1992.
- [6] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer Verlag, 1995.

- [۷] م. ب. منهج، هوش محاسباتی، جلد اول: مبانی شبکه‌های عصبی، انتشارات دانشگاه صنعتی امیرکبیر، تهران، ویرایش اول، ۱۳۷۹.
- [۸] ع. کاشفی کاویانی، س. ع. پورموسوی کانی و ع. جهانبانی اردکانی، "آموزش شبکه‌های عصبی چندلایه با به‌کارگیری الگوریتم PSO"، هشتمین کنفرانس سیستم‌های هوشمند، دانشگاه فردوسی، مشهد، شهریور ۱۳۸۶.

- [9] A. A. Miniani and R. D. Williams, "Acceleration of back-propagation through learning rate and momentum adaptation," in *Proc. Inter. Joint Conf. on Neural Networks*, San Diego, CA, US, vol. 1, pp. 676-679, 1990.
- [10] R. A. Jacobs, "Increased rate of convergence through learning rate adaptation," *Neural Networks*, vol. 1, no. 4, pp. 295-308, 1988.

- [۴۳] ع. راشدی، ح. نظام‌آبادی‌پور و سعید سربزیدی، "الگوریتم جستجوی گرانشی باینری،" هشتمین کنفرانس سیستم‌های هوشمند، دانشگاه فردوسی مشهد، شهریور ۱۳۸۶.
- [۴۴] م. ده‌باشیان و سیدحمید ظهیری، "الگوریتم جستجوی گرانشی نخبه‌گرای پیشرفته،" اولین کنفرانس انرژی‌های تجدیدپذیر و تولید پراکنده ایران، دانشگاه بیرجند، اسفند ۱۳۸۸.
- [۴۵] م. ده‌باشیان و س. ح. ظهیری، "MOGSA: روشی جدید در بهینه‌سازی چندهدفه مبتنی بر الگوریتم جستجوی گرانشی،" شانزدهمین کنفرانس انجمن کامپیوتر ایران، دانشگاه صنعتی شریف، جلد اول، صص ۵۰۷-۵۰۲، اسفند ۱۳۸۹.
- [46] D. E. Rumelhart, E. Hinton, and J. Williams, "Learning internal representation by error propagation," *Parallel Distributed Processing*, vol. 1, 318-362, 1986.
- [47] *UCI Repository of Machine Learning Data Bases*, Department of Information and Computer Science, University of California, Irvine, CA, <http://www.ics.uci.edu/~mllearn/MLRepository.html>.

مریم ده‌باشیان تحصیلات خود را در مقاطع کارشناسی و کارشناسی ارشد مهندسی الکترونیک به ترتیب در سالهای ۱۳۸۵ و ۱۳۸۹ از دانشکده فنی دکتر شریعتی تهران و دانشگاه بیرجند به پایان رسانده است. او هم‌اکنون به عنوان مدرس در دانشگاه‌های غیرانتفاعی مشهد مشغول به تدریس است. زمینه‌های تحقیقاتی مورد علاقه ایشان عبارتند از: تشخیص الگو، الگوریتم‌های بهینه‌سازی ابتکاری و طراحی خودکار مدارهای مجتمع آنالوگ.

سیدحمید ظهیری تحصیلات خود را در مقاطع کارشناسی و کارشناسی ارشد مهندسی برق-الکترونیک به ترتیب در سالهای ۱۳۷۱ و ۱۳۷۴ در دانشگاه صنعتی شریف و دانشگاه تربیت مدرس و دوره دکتری مهندسی برق-الکترونیک را در سال ۱۳۸۴ در دانشگاه فردوسی مشهد به پایان رسانده است. او هم‌اکنون مشغول همکاری با گروه مهندسی الکترونیک و مخابرات دانشکده مهندسی دانشگاه بیرجند می‌باشد. زمینه‌های تحقیقاتی مورد علاقه ایشان عبارتند از: پردازش الگو (تصویر، سیگنال و ...)، تشخیص الگو، روش‌های بهینه‌سازی زیستی و طبیعی، محاسبات نرم و کاربردهای آنها.

- optimization," *Chaos, Solitons and Fractals*, vol. 35, no. 5, pp. 967-979, Mar. 2008.
- [33] 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, no. 2, pp. 1026-1037, Feb. 2007.
- [34] Y. P. Chang and C. N. Ko, "A PSO method with nonlinear time - varying evolution based on neural network for design of optimal harmonic filters," *Expert Systems with Applications*, vol. 36, no. 3, pp. 6809-6816, Apr. 2009.
- [35] R. Pasti and L. N. De Castro, "Bio - inspired and gradient - based algorithms to train MLPs: the influence of diversity," *Information Sciences*, vol. 179, no. 10, pp. 1441-1453, Apr. 2009.
- [36] A. E. Zade Sherme and R. Ghaderi, "An intelligent system for classification of the communication formats using PSO," *Informatica*, vol. 32, no. 2, pp. 213-218, Jul. 2008.
- [37] Y. Marinakis, M. Marinaki, M. Doumpos, and C. Zopoundis, "Ant colony and particle swarm optimization for financial classification problems," *Expert Systems with Applications*, vol. 36, no. 7, pp. 10604-10611, Sep. 2009.
- [38] E. Rashedi, H. Nezamabadi - pour, and S. Saryazdi, "GSA: a gravitational search algorithm," *Information Sciences*, vol. 179, no. 13, pp. 2232-2248, Mar. 2009.
- [۳۹] ع. راشدی و ح. نظام‌آبادی‌پور، "طراحی فیلترهای IIR به وسیله الگوریتم جستجوی گرانشی،" شانزدهمین کنفرانس مهندسی برق ایران، دانشگاه تربیت مدرس، جلد اول، صص. ۴۷۴-۴۶۹، تهران، اردیبهشت ۱۳۸۷.
- [۴۰] ع. راشدی و ح. نظام‌آبادی‌پور، "انتخاب ویژگی با استفاده از الگوریتم جستجوی گرانشی،" سومین کنفرانس فناوری اطلاعات و دانش، دانشگاه فردوسی، مشهد، آذر ۱۳۸۶.
- [41] E. Rashedi, H. Nezamabadi - pour, and S. Saryazdi, "Allocation of static var compensator using gravitational search algorithm," *First Joint Congress on Fuzzy and Intelligent Systems Ferdowsi University of Mashhad, Iran*, Aug. 2007.
- [۴۲] م. ده‌باشیان، سیدحمید ظهیری و ن. مهرشاد، "آموزش شبکه عصبی MLP در فشرده‌سازی تصاویر با استفاده از روش GSA،" ششمین کنفرانس مائسین بینایی و پردازش تصویر ایران، دانشگاه اصفهان، جلد اول، صص. ۴۷۶-۴۷۰، آبان ۱۳۸۹.