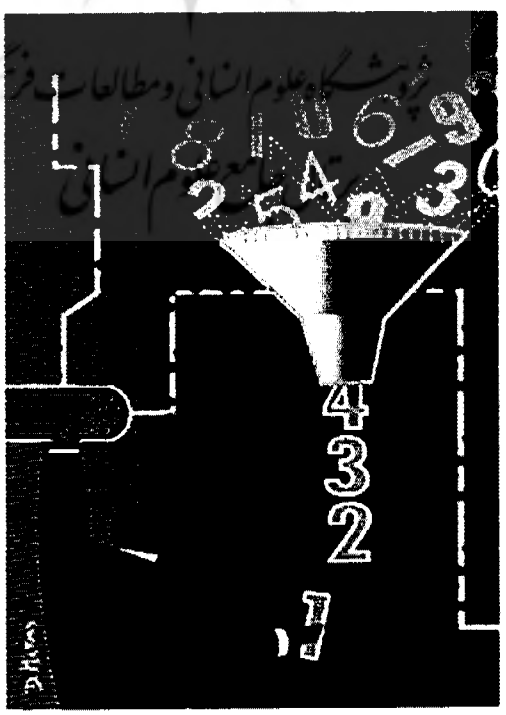
نقش الگوریتم ژنتیک در آموزش و کاهش پیچیدگی شبکه عصبی

بیژن جوهری و حسین فرقانی
(فوق لیسانس حسابداری، عضو هیات علمی دانشگاه ایلام)

شبکه عصبی مصنوعی، ابزار قدرتمندی برای تجزیه و تحلیل داده‌ها و الگوسازی روابط غیر خطی به حساب می‌آید که استفاده از آن طی سال‌های اخیر در حسابداری مالی گسترش یافته است. اغلب مسائل تصمیم‌گیری با استفاده از روش‌های کلاسیک قابل حل هستند، اما به دلیل ساختار ویژه این مسایل و ارتباط تنگاتنگ با افراد تصمیم گیرنده، استفاده از شبکه‌های عصبی برای برآورد ساختار ترجیحات تصمیم گیرنده، به دلیل کارایی قابل ملاحظه‌ای که در حل این نوع مسائل دارد، بیشتر شده است، البته، در اکثر موارد با کمبود اطلاعات (داده) روبه رو هستیم و شبکه عصبی برای ارائه کارایی مناسب به داده‌ها زیاد نیاز دارد؛ به همین دلیل، در این مقاله به نقش کاربرد الگوریتم ژنتیک در شبکه عصبی در برطرف کردن این مشکل پرداخته می‌شود.

۱. مقدمه

علم به طور کلی و فیزیک به طور خاص پیشرفت‌های بعدی خود را مرهون زیست‌شناسی هستند. مشاهده نظام‌های زیستی از یک سو و مشاهده پدیده‌های فیزیکی از سوی دیگر چنین پیشرفتی را ممکن می‌کند. به طوری که در پایان این دو یکی خواهند شد و در پرتو این وحدت حقایق جدیدی روشن خواهد شد که برای هر دو گروه پوشانده مانده است. (ماخ، ۱۹۱۴)



از نظر محققان شبکه‌های عصبی جایگزین روش‌های پردازش معمولی نیست بلکه در کنار آن‌هابرای طرح نظام‌های پیشرفته‌تر به کار می‌آید. هر جا که نیاز به ارتباط پردازش گر با دنیای واقعی و فیزیکی پیش می‌آید نیاز به پردازش الگوها هم وجود دارد و رایانه‌ها در انجام آن ضعیف هستند. برخی محققین شبکه‌های عصبی مصنوعی را نظام‌های حسابگری تعریف می‌کنند که از تعداد زیادی عناصر پردازشگر ساده تشکیل شده است که به صورت پیچیده و با تراکم زیاد با هم در ارتباط هستند. بعضی دیگر از محققین نیز شبکه‌های عصبی مصنوعی را الگوریتم‌هایی با یک سری محدودیت‌ها و خصوصیات ویژه ـ از جمله پردازش موازی اطلاعات ـ تعریف

می‌کنند. (بورک، ۲۰۰۰)

در دهه‌های گذشته پیشرفت‌های چشمگیری در زمینه شبکه‌های عصبی حاصل شده است. هدف اصلی از بکارگیری شبکه‌های عصبی مصنوعی پردازش اطلاعات است. روش پردازش در این شبکه‌ها الهام گرفته از عملکرد مغز انسان است. در واقع هدف تحقیقات اخیر بهره برداری از ساختار شبکه‌های عصبی طبیعی در ایجاد نظام پردازشی است که خصوصیاتی شبیه به مغز انسان داشته باشد. (منهاج، ۱۳۷۹)

بررسی و بهبود ارائه پذیری و توان عمومیت دهی اصلی یک شبکه عصبی اهمیت زیادی دارد. ارائه پذیری و توان عمومیت دهی یک شبکه عصبی به مفهوم دقت و عملکرد آن در نمونه‌های غیر آموزشی است. از مشکلات اساسی که موجب کاهش ارائه پذیری شبکه می‌شود، کمی داده‌های آموزشی و در نتیجه تناسب بیش از حد شبکه نسبت به نمونه‌های آموزشی است؛ در این حالت، به هنگام آموزش مقدار خطا کاهش زیادی می‌یابد اما زمانی که نمونه‌های جدیدی به شبکه ارائه شود، خطای آن غیر قابل قبول می‌شود؛ در صورت اخیر، شبکه تنها نمونه‌های آموزشی را به خاطر می‌سپارد و دیگر برای سایر نمونه‌ها قابل استفاده نیست. از دلایل اصلی تناسب بیش از حد شبکه بر روی نمونه‌های آموزشی وجود اتصالات زاید و پیچیدگی زیاد شبکه است. اتصالات زاید به آن دسته از اتصالاتی گفته می‌شود که در حین آموزش شبکه، به دلیل وجود داده‌های غلط و ناسازگار در نمونه‌های آموزشی و یا به طور تصادفی و برای کاهش بیش از حد خطا، مقادیری به آن اختصاص داده می‌شود؛ این مساله موجب پیچیدگی بیش از حد شبکه و در نتیجه تمرکز بر روی داده‌های آموزشی می‌شود. (سکستون و دیگران، ۲۰۰۲) بنابراین در این مقاله ما به ارائه الگوریتم ژنیتک خواهیم پرداخت که برای تکمیل فرآیند آموزشی و بهبود و افزایش ارائه‌پذیری شبکه عصبی از سوی تصمیم‌گیری‌ها مواجه با کمبود داده‌ها می‌تواند سودمند باشد. دراین الگوریتم با اعمال محدودیتی به طور یکنواخت و با فرض غیرنزولی بودن شبکه نسبت به ورودی‌های آن و در نتیجه کوچکتر بودن فضای جست جو، به تنظیم دوباره پارامترهای شبکه‌هایی خواهیم پرداخت که با الگوریتم گرادیان مزدوج آموزش یافته‌اند. این رویکرد موجب حذف گره‌ها و اتصالات زاید در شبکه و عدم یادگیری داده‌های غلط و ناسازگار در نمونه‌های آموزشی خواهد شد. همچنین در نسل‌های پایانی الگوریتم با در نظر گرفتن یک مقدار جریمه‌ای برای اتصالات غیر صفر، شبکه‌ای از جمعیت نهایی انتخاب می‌شود که کمترین اتصالات و گره‌ها را داشته باشد و در عین حال خطای آن در حد قابل قبول نسبت به سایر شبکه‌ها باشد.

۲.الگوریتم ژنتیک

الگوریتم ژنتیک که هالند[۱] و همکارانش در دانشگاه میشیگان برای اولین بار آن را مطرح کردند، یکی از مهمترین الگوریتم‌های ابتکاری است که از آن برای بهینه‌سازی توابع مختلف استفاده می‌شود. در این الگوریتم اطلاعات گذشته با توجه به خاصیت موروثی بودن الگوریتم، استخراج می‌شود و در روند جست و جو مورد استفاده قرار می‌گیرد. ایده اصلی این روش، علم ژنتیک طبیعی است؛ و دلیل اصلی جذابیت آن نیز این است که با فرض‌های محدود کننده‌ای مثل پیوستگی فضای جست و جو، مشتق‌پذیر بودن، تک ماکزیمی بودن و غیره مواجه نمی‌شود. (سکستون و دیگران، ۲۰۰۲)

۳.تفاوت‌ها و مزایای الگوریتم ژنتیک

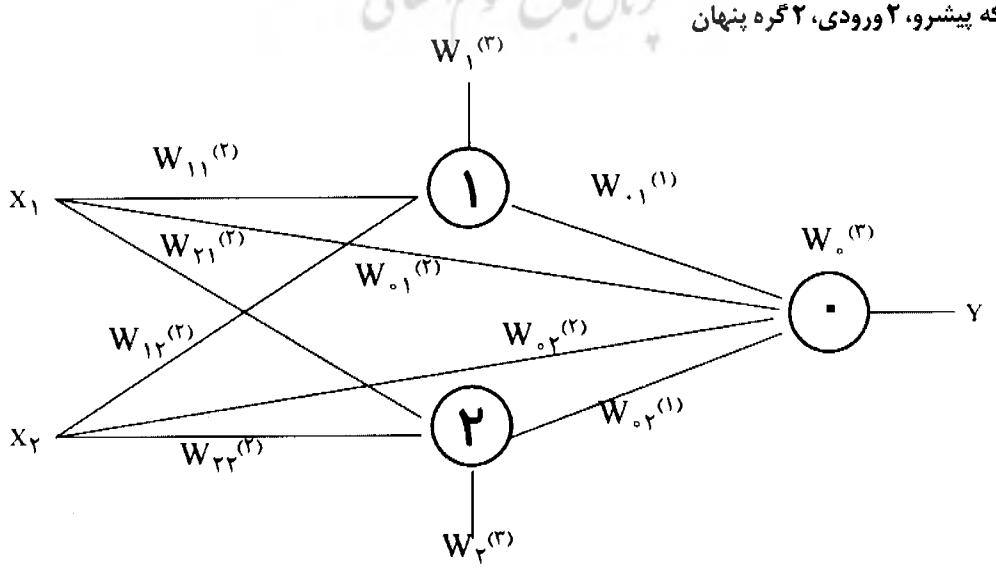
الگوریتم ژنتیک تفاوت‌های اساسی با دیگر روش‌های جست و جو و بهینه‌سازی دارد که می‌توان به شرح زیر به آن‌ها اشاره کرد(همان):

۱.الگوریتم ژنتیک در یک جمعیت از جواب‌ها و با مجموعه‌ای از آن‌ها شروع به جست و جو می‌کند نه با یک جواب؛ و

۲.الگوریتم ژنتیک از قواعد احتمالی استفاده می‌کند نه قواعد طبیعی.

در بسیاری از روش‌های بهینه سازی، با استفاده از یک قانون معین از یک نقطه خاص در فضای جست و جو به نقطه دیگری می‌رویم. این روش نقطه به نقطه از این جهت خطرناک است که فضای جست و جو چند ماکزیممی، فرمان هایی اشتباه برای قرار گرفتن در ماکزیم‌های

نمودار (۱).شبکه پیشرو، ۲ ورودی، ۲ گره پنهان



موضعی داده می‌شود.

در مقایسه با این روش‌ها، الگوریتم ژنتیک به طور همزمان با مجموعه‌ای از نقاط کار می‌کند و بطور موازی از ماکزیمم‌های مختلف بالا می‌رود؛ بنابراین احتمال رسیدن به یک ماکزیمم موضعی کاهش می‌یابد. برخلاف روش‌های دیگر، الگوریتم ژنتیک از قوانین احتمالی در هدایت عمل جست و جو استفاده می‌کند. اما این به معنای یک جست و جوی تصادفی صرف نیست؛ بلکه الگوریتم ژنیتک از انتخاب تصادفی به عنوان ابزاری برای هدایت عمل جست و جو در مناطقی از فضا با احتمال بهبودی بیشتر استفاده می‌کند.

۳.الگوریتم ژنتیک از تابع برازش استفاده می‌کند نه مشتق و گرادیان.

با توجه به مطالب آورده شده می‌توان مزایای مهم الگوریتم ژنتیک را به شرح زیر بیان کرد:

۱) الگوریتم ژنتیک بر خلاف روش‌های دیگر که بیشتر در یک ناحیه به دنبال جواب می‌گردند به طور همزمان چندین ناحیه از فضای جواب‌های شدنی را مورد کاوش قرار می‌دهد. این خصوصیت موجب می‌شود که الگوریتم کمتر در دام نقاط بهینه موضعی گرفتار شود؛

۲) الگوریتم ژنتیک به دلیل طبیعت تکاملی خود، جواب‌ها را بدون توجه به طرز کار ویژه مساله، جست و جو می‌کند و می‌تواند با هر نوع تابع هدف و محدودیتی (خطی و غیرخطی) در هر فضای جست و جو (گسسته، پیوسته یا مرکب) کار کند؛

۳) الگوریتم ژنتیک انعطاف‌پذیری بالاتری دارد؛

۴) عملکردهای ژنتیکی، الگوریتم ژنتیک را در اجرای جست و جوی سراسری، بسیار کارآ می‌کنند؛ و

۵) در این روش محاسبات به طور دقیق انجام شده و هیچگونه تقریبی نظیر خطی‌سازی تابع هدف، گرد کردن نتایج و تغییر متغیرهای گسسته به پیوسته و بالعکس وجود ندارد. همچنین این روش بسیار ساده است.

۴.مفاهیم اساسی، الگوریتم ژنتیک

در الگوریتم ژنتیک قبل از هر چیز ساز و کاری برای نمایش هر جواب مساله به صورت یک کروموزوم تعریف می‌شود. سپس مجموعه‌ای از کروموزم‌ها که در حقیقت بیانگر مجموعه‌ای از جواب‌های مساله هستند به عنوان جمعیت اولیه در نظر گرفته می‌شود. بعد از این مرحله، باید با بکارگیری عملکردهای ژنتیک اقدام به تولید کروموزم‌های جدید موسوم به نوزاد کرد. پس از تولید تعدادی کروموزم‌های منتخب برابر اندازه جمعیت اولیه، فرایند انتخاب، مبتنی بر مقدار برازندگی هر رشته است؛ در هر رشته‌ای اغلب تابع برازش برابر با همان تابع هدف مساله بهینه سازی در نظر گرفته می‌شود. تاکنون یک بار تکرار یا یک نسل از الگوریتم تولید شده است. الگوریتم، بعد از تولید چندین نسل به تدریج به سمت جواب بهینه همگرا می‌شود. شرط توقیف مساله می‌تواند طی کردن تعداد معینی تکرار باشد که از قبل توسط کاربر تعیین شده است؛ یا عدم تغییر جواب نهایی الگوریتم در چند تکرار مشخص از آن و یا شرط خاص دیگری باشد. (هرا و دیگران، ۱۹۹۶)

۵.الگوریتم ژنتیک در شبکه عصبی

در این قسمت ضمن ارائه اجزای الگوریتم ژنتیک در شبکه عصبی، به ارزیابی و تشریح آن می‌پردازیم. ترتیب ارائه این اجزا بدین صورت است: نحوه کد کردن و نمایش کروموزوم، چگونگی ایجاد جمعیت اولیه، عملگر انتخاب، تعیین تابع برازش، عملگرهای ژنتیکی (جهشی و تقاطعی).

برای سادگی توضیح هر کدام از موارد بالا، شبکه زیر را با دو گروه در لایه پنهان و دو ورودی در نظر می‌گیریم.

۵.۱.کد کردن و نمایش کروموزوم

از آنجا که ساختار شبکه عصبی در سرتاسر الگوریتم تغییر نمی‌کند و توابع فعال سازی گره‌ها یکسان است، هر کروموزوم به صورت رشته‌ای از وزن‌ها و حدود آستانه‌ای گره‌ها نمایش داده می‌شود. بنابراین، هر کروموزوم نشانگر تمام پارامترهای قابل تنظیم شبکه عصبی است (آرنوان و دیگران، ۱۹۹۶)، مقادیر ژن‌ها (وزن‌ها و حدود آستانه‌ای) در هر کروموزوم به دلیل برتری کدهای حقیقی نسبت به کدهای دو دویی در الگوریتم ژنتیک، به صورت کدهای حقیقی در نظر گرفته شده است. از مزیت‌های نمایش ژن‌ها به صورت کدهای حقیقی نسبت به کدهای دو دویی، می‌توان به موارد زیر اشاره کرد (همان):

۱) نیازی به تبدیل کدها به مقادیر حقیقی برای محاسبه تابع برازش الگوریتم نیست. همچنین طول کروموزم‌ها به طور قابل ملاحظه‌ای کاهش می‌یابد که این امر تاثیر زیادی در فراوانی بکارگیری عملکردهای الگوریتم دارند؛

۲) مشکل تقریب و افزایش دقت محاسبه وجود ندارد؛

۳) می‌توان عملکردهای جهشی و تقاطعی و متنوعی را جهت افزایش کارایی و اثربخشی الگوریتم بکار گرفت؛ و

۴) الگوریتم از یک نسل به نسل دیگر نتایج سازگارتری ارائه می‌دهد.

موضوع دیگر در نمایش کروموزم‌ها، نحوه کنار هم قرار گرفتن ژن‌هاست. در اینجا به دلیل این که وزن‌ها و حدود آستانه‌ای گره‌ها به همدیگر وابسته است، تمام اتصالات یک گره در کنار هم و گره‌ها به صورت لایه به لایه نمایش داده می‌شود (آرنوان و دیگران، ۱۹۹۶؛ مودتانا، ۱۹۹۵). برای نمونه کروموزوم مربوط به نمودار (۱) در زیر آورده شده است.

$$\frac{W_{0}^{(2)}, W_{10}^{(2)}, W_{20}^{(2)}}{\sqrt{3}}, \frac{W_{0}^{(1)}, W_{21}^{(2)}, W_{22}^{(2)}, W_{2}^{(3)}}{\sqrt{2}}, \frac{W_{0}^{(1)}, W_{11}^{(2)}, W_{12}^{(2)}, W_{1}^{(3)}}{\sqrt{1}}$$

۵.۲.جمعیت اولیه

برای تولید جمعیت اولیه از روش گرادیان مزدوج استفاده می‌شود و هر کدام از افراد (کروموزوم‌ها) آن شامل وزن‌ها و حدود آستانه‌ای است که با این روش تنظیم شده‌اند. دلیل بکارگیری این روش در تولید جمعیت اولیه، پیچیدگی فضای جست و جو، نامحدود بودن دامنه متغیرها و ناتوانی روش تصادفی است. همچنین در روش گرادیان مزدوج به دلیل استفاده

از مشتق و گرادیان، دستیابی به مجموعه گره‌هایی که اتصالات آن‌ها رفتار خوبی در عملکر کلی شبکه ایجاد می‌کند، امکان‌پذیر است. این مساله موجب بوجود آمدن گونه‌های خوب در جمعیت اولیه و همگرایی سریع و مناسب الگوریتم می‌شود. (استر مارک، ۱۹۹۹)

۵.۳. عملگر انتخاب

انتخاب کروموزوم‌ها به صورت انتخاب مسابقه انجام می‌شود؛ اندازه مسابقه ۳ در نظر گرفته شده است. در این روش، از فضای نمونه‌گیری به تعداد جمعیت اولیه، مجموعه ۳ عضوی انتخاب می‌شود و از هر مجموعه بهترین عضو انتخاب می‌گردد. دلیل در نظر گرفتن اندازه ۳ برای مسابقه، فضای نمونه‌گیری است که به صورتی توسعه یافته است. (پندارکر و راجر، ۲۰۰۲)

پس از اتمام عمل انتخاب در یک نسل، برای جلوگیری از عدم همگرایی الگوریتم در مراحل پایانی، بدترین فرد جمعیت انتخابی با بهترین فرد که تاکنون پیدا شده، جایگزین می‌شود. (دورسی و می یر، ۱۹۹۵)

۵.۴. تابع برازش

تابع برازش الگوریتم از سه قسمت اصلی تشکیل شده است که در ادامه هر کدام به طور مجزا توضیح داده می‌شود.

الف. خطای شبکه عصبی (RMSE)

برابر ریشه دوم میانگین مجموع مربعات اختلاف خروجی شبکه با مقادیر متناظر است. (سکستون و دیگران، ۲۰۰۲؛ سکستون و دیگران، ۱۹۹۸)

$$RMSE=\sqrt{\frac{\sum_{P=1}^{P}(Y^{P}-Z^{P})}{P}}$$

ب. شرط بطور یکنواخت غیرنزولی بودن خروجی شبکه عصبی

این شرط به صورت یک مقدار جریمه‌ای در تابع ارزش آورده شده است؛ این مقدار برای شبکه عصبی شکل (۱) به صورت زیر خواهد بود:

$$D=-(g_1 W_{11}^{(1)} W_{11}^{(2)} + g_2 W_{21}^{(1)} W_{21}^{(2)} + W_{11}^{(2)} + g_1 W_{12}^{(1)} W_{12}^{(2)} + g_2 W_{22}^{(1)} W_{22}^{(2)} + W_{12}^{(2)})$$

$$g_1 = \begin{cases} 0 & (W_{11}^{(1)} W_{11}^{(2)}) \geq 0 \\ 0.25 & (W_{11}^{(1)} W_{11}^{(2)}) < 0 \end{cases} \quad , \quad g_2 = \begin{cases} 0 & (W_{12}^{(1)} W_{12}^{(2)}) \geq 0 \\ 0.25 & (W_{12}^{(1)} W_{12}^{(2)}) < 0 \end{cases} \quad , \quad j=1,2$$

پ. شرط حذف گره‌ها و اتصلات زاید

این شرط به صورت یک مقدار جریمه در تابع برازش آورده شده و شامل دو قسمت است (nozer) × (rmse)

nozer نشانگر تعداد ژن‌های مخالف صفر در یک کروموزوم است. این مقدار جریمه‌ای در مراحل نهایی الگوریتم، اتصالات و گره‌های زایدی را که تاثیر چندانی در عملکرد کلی شبکه ندارد، حذف می‌کند و موجب کاهش پیچیدگی شبکه می‌شود.

بعد از تشریح اجزای تابع برازش، نوبت به ارائه فرم کلی و ارزیابی پارامترهای آن می‌رسد. فرم کلی تابع برازش به صورت زیر است:

$$\text{Min : Fit - Rmse} + (\text{monoDec})^{pf1} + (\text{nozer})^{pf2} \times (\text{Rmst}) \times \text{sign}(1 + pf3)$$

در عبارت بالا sign، تابع علامت و t نشانگر شماره تکرار الگوریتم است pf۳pf۲pf۱. پارامترهای عبارات جریمه‌ای است و مقادیر آن‌ها با توجه به روند در نظر گرفته شده برای تابع برازش، تنظیم می‌شود.

۵.۵. عملگرهای ژنتیکی

الف - عملگر جهشی

در الگوریتم ما از سه عملگر جهشی استفاده می‌کنیم. عملگر اول برای ایجاد گونه‌های جدید و دو عملگر دیگر برای حذف اتصالات زاید در شبکه بکار گرفته می‌شود. در زیر نحوه انجام هر کدام از این عملکرها توضیح داده می‌شود.

عملگر گره‌ای[۲] (NNN): در این عملگر ابتدا یکی از گره‌های شبکه در صورت تصادفی (و بطور یکنواخت) انتخاب می‌شود و سپس به هرکدام از اتصالات آن یک مقدار تصادفی تولید شده از توزیع نرمال[۳] با میانگین و انحراف معیار ۱ اضافه می‌شود. (آرنوان و دیگران، ۱۹۹۶؛ مودتانا، ۱۹۹۵)

عملگر صفر[۴] (Nzm): این عملگر برای حذف اتصالات زاید در شبکه بکار گرفته می‌شود؛ نحوه اعمال آن بدین صورت است که ابتدا یکی از اتصالات شبکه به صورت تصادفی انتخاب و مقدار آن برابر صفر قرار داده می‌شود. (سکستون و دیگران، ۲۰۰۲)

عملگر صفر گره‌ای[۵] (NNZM): این عملگر برای حذف گره‌های

زاید در شبکه در نظر گرفته می‌شود؛ نحوه اعمال آن بدین صورت است که در آن ابتدا به صورت تصادفی گره‌ای از شبکه انتخاب می‌شود و به ازای هر کدام از اتصالات آن به طور تصادفی عدد صفر یا یک تولید می‌شود. در صورتی که عدد تولید شده برابر یک باشد، مقدار آن اتصال برابر صفر قرار داده می‌شود.

ب - عملگر تقاطعی

در الگوریتم ژنتیک عملگر تقاطعی نقش بسیار مهمی ایفا می‌کند و در واقع توان آن تاثیر به سزایی در نتیجه نهایی الگوریتم دارد. این عملگر تنها ابزاری است که موجب انتقال اطلاعات بین کروموزوم یک جمعیت می‌شود و در حالت کلی می‌توان عملگر آن را از دید و فاکتور زیر مورد بررسی قرار داد. (آرنوان و دیگران، ۱۹۹۶)

الف - انتقال اطلاعات مفید از چندین کروموزوم به یک راه حل بهتر و تکرار این روند در طول الگوریتم و یا به عبارت دیگر افزایش تعداد ژن‌های مطلوب در کروموزوم‌ها و تثبیت مکان این ژن‌ها.

ب - برعکس حالات بالا یعنی کاهش و پراکندن ژن‌های مطلوب و مفید مکان‌های تثبیت شده در کروموزوم‌ها.

این دو عامل برای هر عملگر تقاطعی در الگوریتم ژنتیک مطرح است: عملگر تقاطعی مناسب خواهد بود که تعدیل بین این دو حالت در آن وجود داشته باشد. در استفاده از الگوریتم ژنتیک برای شبکه‌های عصبی به دلیل وابستگی شدیدی که بین ژن‌های یک گروه وجود دارد، بکارگیری عملگرهای معمول (در الگوریتم ژنتیک با کدهای حقیقی) موجب انتقال مناسب اطلاعات بین کروموزوم‌ها نمی‌شود و تعدیل مطلوبی برای دو عامل مذکور ایجاد نمی‌شود. وابستگی بین وزن‌های یک گره در آموزش شبکه را می‌توان بدین صورت بیان کرد که تنها تنظیم یکی از اتصالات در آن (یک گره) تاثیر بسیار ناچیز و یا شاید هیچ تاثیری در عملکرد کلی شبکه ندارد؛ زمانی که به اتصالات یک گره به طور همزمان مقادیر مناسبی تخصیص داده می‌شود، خطای شبکه تا حد قابل توجهی کاهش می‌یابد. از طرف دیگر به دلیل روابطی که بین اتصالات یک گره در شبکه وجود دارد، احتمال اینکه بعضی از وزن‌های آن از یک کروموزوم و بقیه از کروموزوم دیگر موجب بهبود عملکرد شبکه شود، بسیار کم است. البته این احتمال تا حدی با نمایش کروموزوم‌ها به صورت گره‌ای افزایش داده شده است. (آرنوان و دیگران، ۱۹۹۶؛ مودتانا، ۱۹۹۵)

نتیجه گیری

کاربرد الگوریتم ژنتیک در آموزش شبکه عصبی می‌تواند ابزار بسیار توانمندی برای تصمیم‌گیری فراهم کند. نتیجه تصمیم‌گیری می‌تواند تا آن اندازه قوی باشد که داده‌های کم به تصمیم‌گیری بهینه منجر شود، زیرا شبکه عصبی در مواردی که داده‌های ورودی زیاد باشد کاربرد مناسب دارد و با داده‌های کم جهت آموزش مقدار خطا بسیار کاهش می‌یابد اما زمانی که نمونه‌های جدیدی به شبکه ارائه می‌شود، خطای آن غیرقابل قبول می‌شود که در این صورت شبکه تنها نمونه‌های آموزشی را به خاطر می‌سپارد و قابل استفاده برای نمونه‌های دیگر نیست.

پی نوشت‌ها:

1. G. Holland- 1975.
2. Neural Network node mutation
3. normal disribution
4. Neural Network Zero mutation
5. Neural Network Node Zero mutation

منابع:

1- Minghe Multiple objective programming using Tchebycheff programs and Artificial Neural Network; COMP.OPS.RES.VOL2. pp 601 - 620 , 2000.

2- Hakin.S ; Neural networks a comprehensive foundation; macmillan college publishing, 1994.

۳. منهاج م.ب: مبانی شبکه‌های عصبی، هوش محاسباتی: انتشارات دانشگاه صنعتی امیرکبیر، ۱۳۷۹.

4- Burke I.I and ignizioj.p; neural networks in Business Technirues and application for the operations researcher; COMP . OPS. RES. Vol 27, pp 1023 - 1044 , 2000.

5- Gen, Mand chang. R; Cenetic Algorithm and Engineering optimization; John wiley, 2000.

6- Herrea . F and Lozano.m and Verdegay j.l; Tackling real-coded genetic Algorithm; DGICYTPB, 1996.

7- Dorsey R.E and Mayer w.g; Genetic Algorithms for Estimation problems with multiple ptima, Nondifferentiability o and other irregular features; journal of Business and Economic statistics; vol 13 no 1, 1995.

8- Ostermark R;solving Lrregular Econometric, and Mathematical optimization Hybrid Algorithm; Comutational Economics problems with a Genetic vol 13, pp 103 - 115 , 1999.

9- Arnovan.R and lakhmi J. and Rayj; neural network training using Genetic Algorithm; wo rld scientific publishing; 1996.

10- Modtana D.J ; neural network weight selection using Genetic Algoritm intelligent Hybrid systems; john wiley, 1995.

11- Sextoon R.S and Dorsey R.E and sikander N.A, simultaneous optimization of neural Network Function and architecture algorihm; Decision support systems, 1034, 2002.

12- Pendharker P.c and Roddger J.A/ An Empirical study of Impact of crossver operators on the performance of Non-Binury Genetic Algorithm Bosed Neural Approaches for Classification; Computers and operations Research, 2002

13- Sexton R.S and Dorsey R.E and jehnson j.D; Toward global optimization of Neural networks, A comparison of the Genetic Algorithm and Backpropagation; Decision support systems vol 22, pp 171-185, 1998.