

SID



ابزارهای
پژوهش



سرویس ترجمه
تخصصی



کارگاه های
آموزشی



بلاگ
مرکز اطلاعات علمی



سامانه ویراستاری
STES



فیلم های
آموزشی

کارگاه های آموزشی مرکز اطلاعات علمی



آموزش مهارت های کاربردی در تدوین و چاپ مقالات ISI

آموزش مهارت های کاربردی
در تدوین و چاپ مقالات ISI



روش تحقیق کمی

روش تحقیق کمی



آموزش نرم افزار Word برای پژوهشگران

آموزش نرم افزار Word
برای پژوهشگران

مقایسه الگوریتم ژنتیک با الگوریتم های پس انتشار خطا در آموزش شبکه های عصبی برای شناسایی و مدل سازی دینامیک غیرخطی هواپیما

علیرضا رودباری^۱، فریبرز ثقفی^۲

۱ و ۲- دانشگاه صنعتی شریف دانشکده هوافضا

چکیده

تاکنون تحقیقات فراوانی در مورد کاربردهای مختلف شبکه های عصبی صورت گرفته است. سابقه بیش از دو دهه استفاده گسترده از شبکه های عصبی مصنوعی در کاربردهای متنوع از جمله بهینه سازی، مدل سازی و شناسایی سیستم و همچنین توانایی تقریب زنی، یادگیری و تعمیم دهی سبب شده است که این شبکه ها جایگزین مناسبی برای روش های کلاسیک شناسایی و مدل سازی سیستم های غیرخطی از جمله هواپیما باشند. علیرغم اینکه شبکه های عصبی در علوم مختلف از موفقیت قابل توجهی برخوردار بوده اند؛ اما تعیین الگوریتم آموزش و معماری شبکه برای داشتن تعمیم دهی مناسب در کاربردهای مدل سازی هنوز به عنوان یک مشکل اساسی باقی مانده است. یکی از مواردی که شبکه های عصبی را در این کاربردها محدود نموده است عدم توانایی یادگیری و بدست آوردن وزن های مناسب در شبکه عصبی برای سیستم های غیرخطی می باشد. اغلب روش های آموزش شبکه های عصبی بر پایه الگوریتم های گرادینتی و نوعاً گونه ای از پس انتشار خطا برای یافتن پارامترهای وزنی شبکه می باشد. محدودیت های استفاده از جستجوی گرادینتی در مسائل بهینه سازی پیچیده ای مانند آموزش شبکه های عصبی مصنوعی موجب عملکرد متناقض و غیرقابل پیش بینی آن می گردد. در سال های اخیر استفاده از الگوریتم های ژنتیک، که برای غلبه بر چنین مشکلاتی به خوبی در یک فضای بزرگ و پیچیده جستجو کاوش می کنند، بسیار رایج شده است. در این مقاله مزایا و خصوصیات الگوریتم های مختلف پس انتشار خطا از قبیل لونبرک، کواری نیوتن و گرادینت مزدوج، و الگوریتم ژنتیک در روند آموزش شبکه عصبی بازگشتی در شناسایی و مدل سازی دینامیک غیرخطی هواپیما برای داشتن تعمیم دهی مناسب مورد بررسی و مقایسه قرار گرفته است.

واژه های کلیدی: شبکه عصبی - الگوریتم ژنتیک - الگوریتم پس انتشار خطا - دینامیک هواپیما - شناسایی و مدل سازی

مقدمه

شبکه های عصبی با توجه به ویژگی های تقریب زنی عمومی و تعمیم دهی، می توانند جانشین مناسبی برای روش های کلاسیک مدل سازی و شبیه سازی اغلب سیستم های دینامیکی از جمله وسایل پرنده باشند [۱-۳]. علیرغم اینکه شبکه های عصبی در علوم مختلف از موفقیت قابل توجهی برخوردار بوده اند؛ اما تعیین الگوریتم آموزش و معماری شبکه برای داشتن تعمیم دهی مناسب در کاربردهای مدل سازی هنوز به عنوان یک مشکل اساسی باقی مانده است. مقصود از تعمیم دهی داشتن عملکرد مناسب به ازای داده هایی که در آموزش شبکه دخالت نداشته اند، است. تعمیم دهی در موضوع شناسایی به صورت آفلاین (بدون داشتن پس خور از سیستم

اصلی) بسیار مهم است. هرچه شبکه عصبی از قدرت تعمیم دهی بالاتری برخوردار باشد، از اعتبار بیشتری در پیش بینی رفتار سیستم برخوردار خواهد بود. افزودن پسخور داخلی باعث افزایش حافظه دینامیکی شبکه می شود و افزایش حافظه دینامیکی یکی از راهکارهای افزایش قدرت تعمیم دهی در مدل سازی و شناسایی سیستم های دینامیکی غیرخطی است [۱-۳]؛ اما در عین حال باعث بزرگ تر شدن شبکه می گردد. بزرگ تر شدن شبکه یادگیری را سرعت می بخشد اما تأثیر منفی بر روی تعمیم دهی خواهد داشت [۴]. کوچک نگاه داشتن شبکه از راه های مختلفی میسر است. از جمله: هرس کردن، الگوریتم کاهش و الگوریتم افزایشی [۵]. این روش ها معمولاً بسیار وقت گیر بوده و از سوی دیگر نیز تضمین قطعی برای بهبود تعمیم دهی نمی دهند [۵]. بنابراین علاوه بر بهینه سازی الگوریتم آموزش شبکه، بهینه سازی معماری نیز می تواند به تعمیم دهی و عملکرد بهتر شبکه، برای سیستم های پیچیده کمک بزرگی نماید. عملکرد بهتر شبکه را می توان به میزان قابل ملاحظه ای به روش ها و الگوریتم های جستجو استفاده شده در حین آموزش نسبت داد. بر اساس مراجع مختلف، اغلب روش های آموزش برای یافتن پارامترهای وزنی شبکه های عصبی مبتنی بر الگوریتم های گرادینتی و نوعاً گونه ای از پس انتشار خطا [۶] هستند. اما محدودیت های استفاده از این الگوریتم ها در مسائل بهینه سازی موجب عملکرد متناقض و غیرقابل پیش بینی می گردد [۷]. همگرایی در روش های آموزش شبکه های عصبی بر پایه الگوریتم های گرادینتی به وضعیت بهینه سراسری تضمین شده نیست و احتمال به دام افتادن در مینیمم محلی زیاد است [۸]. پارامترهای اولیه به طور تصادفی انتخاب می شوند و می توان مشاهده کرد که در اجراهای مختلف برنامه شبکه عصبی، نتایج کاملاً متمایزی در تعداد تکرارهای مساوی الگوریتم حاصل می شود. همچنین آموزش شبکه های عصبی با الگوریتم های جستجوی گرادینتی باعث انطباق بیش از حد شبکه به داده های آموزشی می گردد. انطباق بیش از حد شبکه، باعث کاهش قابلیت تعمیم دهی شبکه می شود. انطباق بیش از حد نیز هنگامی رخ می دهد که با وجود اینکه شبکه در نگاشت داده های آموزشی به خروجی مطلوب بسیار دقیق است، اما به خاطر فرا نگرفتن الگوی حاکم در پاسخ به ورودی های جدید، ناموفق عمل می کند [۹]. بر طبق گزارش محققین، در مقایسه با دیگر روش های یادگیری، الگوریتم های ژنتیک می تواند جایگزین روش های پس انتشار در آموزش شبکه های عصبی شود و عملکرد شبکه را بهینه سازد [۱۰]؛ اما روند یادگیری در الگوریتم ژنتیک نیز بسیار کند می باشد؛ بنابراین این نوع الگوریتم هم برای شناسایی آن لاین ممکن است مناسب نباشد. مدل سازی وسایل پرنده در زمره سیستم های دینامیکی غیرخطی پیچیده و کوبله می باشند که به دلیل وجود نقاط مینیم محلی با روش های متداول به سختی می توان شبکه را آموزش داد و در صورت آموزش نیز تعمیم دهی

۱- دانشجوی دکتری هوافضا گرایش دینامیک پرواز و عضو هیأت علمی دانشگاه هوایی شهید ستاری (alirezaroudbari@ae.sharif.edu)

برای بقا داشته باشند زنده می‌مانند و به تولید مثل می‌پردازند. این ویژگی باعث می‌شود الگوریتم ژنتیک قادر به بهینه‌سازی در نقاط سراسری باشد. استفاده از الگوریتم‌های ژنتیک در رابطه با شبکه‌های عصبی به چند روش صورت می‌گیرد: برای یافتن وزن‌های شبکه عصبی، پیدا کردن معماری بهینه، بهینه‌سازی هم‌زمان معماری و وزن‌های شبکه و تحلیل یک شبکه عصبی. در این مطالعه از دو روش الگوریتم ژنتیک برای یافتن وزن‌های شبکه با معماری ثابت، با نام الگوریتم ژنتیک معمولی (NGA)^۱ و دومین روش بهینه‌سازی هم‌زمان معماری و وزن‌های شبکه با نام الگوریتم ژنتیک اصلاح شده (MGA)^۲ استفاده شده است.

برای آموزش شبکه عصبی در روند بهینه‌سازی با الگوریتم ژنتیک معمولی برای پیدا کردن وزن‌های بهینه شبکه عصبی، ارتباط‌های وزنی روبه جلو، بازگشتی، ارتباط‌های خود بازگشتی و بایاس استفاده می‌شوند. هر نسل در الگوریتم ژنتیک جمعیتی از اعضا است و هر عضو با یک کروموزوم که ویژگی‌های آن را به صورت کد شده در بر دارد، مشخص می‌شود. در شبکه‌های عصبی، کروموزوم‌ها آرایه‌های سطری شامل وزن‌ها (بیت‌ها) هستند که شامل تمام وزن‌ها و بایاس‌های شبکه می‌باشند. کافی است تمام وزن‌ها و بایاس‌های شبکه در یک رشته به دنبال هم ذخیره شوند. همین‌طور می‌توان وزن‌های مربوط به ارتباط‌های داخلی و پس‌خور را به کروموزوم‌ها اضافه کرد و در فرایند آموزش دخالت داد. کروموزوم‌ها معمولاً به صورت باینری کد می‌شوند؛ ولی در مسئله شبکه عصبی این کار باعث زیاد شدن طول کروموزوم و در نتیجه درست کار نکردن عملگرهای الگوریتم ژنتیک می‌شود؛ لذا در این تحقیق به منظور اجتناب از افزایش طول کروموزوم، پارامترهای وزنی و بایاس‌ها با اعداد دهگانی بیان شده‌اند. تابع شایستگی یا ارزیابی، نقش محیطی را ایفا می‌کند که راه حل‌های خوب و بد را تشخیص می‌دهد. در هر نسل، راه حل‌های نسبتاً خوب برای یک مشکل تکثیر شده و برای جایگزینی راه حل‌های نسبتاً بد که ماندگار نیستند تولید مثل می‌کنند. از آنجایی که در الگوریتم ژنتیک هدف مینیمم کردن اختلاف خروجی شبکه عصبی و خروجی پاسخ دینامیکی سیستم مورد نظر یعنی پاسخ دینامیکی هواپیماست، معکوس میانگین مربعات خطا (MSE) معیار مناسبی به عنوان تابع شایستگی است.

$$J = MSE = \sum_{i=1}^{n_d} (O_i - T_i)^2 \quad (1)$$

$$Fit = \frac{1}{MSE} \quad (2)$$

در این رابطه، J تابع هزینه، Fit معیار شایستگی T_i خروجی مطلوب (از آزمایش‌های پروازی بدست می‌آید)، O_i خروجی شبکه عصبی و n_d تعداد نقاط، می‌باشد.

معیار شایستگی هر یک از وزن‌ها در جمعیت ژنتیکی به طور معمول شامل میانگین مربعات خطا نرون‌های خروجی روی یک دوره آموزشی است، که معمولاً پس از آنکه شبکه عصبی برای یک دوره زمانی آموزش دید اندازه‌گیری می‌شود. در هر عضو، هرچه خروجی شبکه و خروجی مطلوب در نقاط داده بیشتری به هم نزدیک‌تر شوند، میانگین مربعات خطا کوچک‌تر و شایستگی آن عضو بالاتر می‌رود.

الگوریتم ژنتیک اصلاح شده همانند الگوریتم ژنتیک معمولی است با این تفاوت که در این الگوریتم از یک اپراتور جدید به نام "جهش-۲" استفاده می‌شود که به کمک آن می‌توان به طور تصادفی برخی از وزن‌ها را صفر و از شبکه خارج کرد. در این اپراتور برای هر بیت از هر کروموزوم یک عدد

مناسبی نخواهد داشت [۱ و ۲]؛ بنابراین ارائه الگوریتم آموزش مناسب در روند شناسایی و مدل‌سازی وسایل پرنده از اهمیت بسیار بالایی برخوردار است. در این مقاله مزایا و خصوصیات الگوریتم‌های مختلف پس انتشار خطا از قبیل لونیبرگ، کوازی نیوتن و گرادیان مزدوج، و الگوریتم ژنتیک در روند آموزش شبکه عصبی بازگشتی در شناسایی و مدل‌سازی دینامیک غیرخطی هواپیما برای داشتن تعمیم دهی مناسب مورد بررسی و مقایسه قرار گرفته است. داشتن کمترین خطا تابع هزینه برای داده‌های تعمیم-دهی و یا بیشترین درصد انطباق خروجی شبکه با داده مورد نظر، مهم‌ترین معیار عملکرد مناسب برای هر یک از الگوریتم‌های آموزش می‌باشد.

الگوریتم آموزش در شبکه‌های عصبی

آموزش در شبکه‌های عصبی به این معناست که شبکه اتصالات وزن‌دار خود را طوری تنظیم کند که خروجی شبکه با خروجی مورد نظر یکسان شود. برای آموزش شبکه ابتدا باید وزن‌ها و بایاس‌های شبکه وزن‌دهی اولیه شوند. هنگامی که وزن‌های و بایاس شبکه مقداردهی اولیه شدند، شبکه برای آموزش دیدن، آماده است. شبکه می‌تواند برای تقریب تابع (رگرسیون غیرخطی)، ارتباط الگویی، و یا طبقه بندی الگویی مورد آموزش قرار گیرد.

برای ارزیابی کیفیت مدل، یک معیار عینی یا تابع هزینه که "خوبی" فیت بودن را اندازه‌گیری می‌کند، استفاده می‌شود. تابع هزینه بین مدل و داده‌های اندازه‌گیری شده ارتباط ایجاد می‌کند. برای انتخاب بهترین مدل باید این تابع هزینه مینیمم یا ماکزیمم شود؛ لذا بهینه‌سازی جایگاه خاصی در مباحث شناسایی سیستم دارد. بسیاری از الگوریتم‌ها و روش‌های یادگیری در شناسایی سیستم، مبتنی بر کمینه کردن یک تابع می‌باشد و توابع نیز معمولاً از نوع انرژی و یا خطا می‌باشند؛ بنابراین برای بدست آوردن ساختار بهینه مدل و قوانین آموزش و مقادیر پارامترهای بهینه، می‌توان از روش‌های بهینه‌سازی جهت آموزش مدل استفاده نمود. بنابراین؛ بسیاری از الگوریتم‌های بهینه‌سازی به طور گسترده‌ای به منظور آموزش شبکه‌های عصبی برای حل مسائل پیچیده غیر خطی طراحی شده و توسعه یافته‌اند. به طور کلی الگوریتم‌های آموزش به دو بخش الگوریتم‌های تکاملی و گرادیانی و نوعاً گونه‌ای از نوع پس‌انتشار خطا تقسیم بندی می‌شوند.

الگوریتم‌های تکاملی

الگوریتم ژنتیک در میان الگوریتم‌های تکاملی بیشتر مورد توجه و استفاده قرار گرفته است. الگوریتم ژنتیک متعلق به کلاسی از الگوریتم‌های جستجوی تصادفی بر پایه جمعیت می‌باشد که الهام گرفته از فرضیه تکامل تدریجی است که الگوریتم‌های تکاملی نامیده می‌شوند.

یک الگوریتم ژنتیک با جمعیتی از راه حل‌ها، کروموزوم‌ها و فرایند تصادفی آغاز می‌شود و با به‌کارگیری عملگرهای الهام گرفته شده از فرآیندهای موجود در طبیعت به راه حل‌های بهتری دست می‌یابد. عملگرهای مورد استفاده در این الگوریتم عبارتند از: ترکیب، جهش، نخبه‌گزینی و انتخاب تصادفی از میان اعضا صورت می‌گیرد. تشکیل جمعیت اولیه به صورت تصادفی است و اپراتورهایی که نسل‌های جدید را می‌سازند دارای عناصر تصادفی هستند. اعضای جمعیت در هر نسل با هم به رقابت می‌پردازند و آن‌هایی که بر حسب خصوصیات برترشان، قدرت و شایستگی بیشتری

¹ - Normal genetic algorithm

² - Modified genetic algorithm

روش‌هایی مانند گرادیان مزدوج، لوبنرک-مارکوآرت (LM) ^۳ و کوآزی-نیوتن از نوع الگوریتم‌های سریع می‌باشند.

روش شیب تند^۴

این روش با نام روش گرادیان شناخته شده است و از روش‌های موثر و ساده جهت کمینه کردن توابع غیرخطی در فضای پارامتری چند بعدی می‌باشد. این روش اساس روش‌های بهینه‌ساز محدودیت‌دار و بدون محدودیت را فراهم می‌کند. در رابطه (۵) اگر ماتریس G را یک ماتریس همانی در نظر بگیریم در این صورت روش بیشترین کاهش یا روش نزول تند بدست می‌آید:

$$X(k+1) = X(k) - \alpha(k)g(x(k)) \quad (6)$$

$g(x(k))$ بردار گرادیان تابع هزینه می‌باشد و جستجو در جهت مخالف آن صورت می‌گیرد.

روش نیوتن نیز همانند روش شیب تند می‌باشد با این تفاوت که بردار جهت کاهش با استفاده از مشتق دوم تابع هزینه بدست می‌آید. در رابطه کلی (۵) اگر $\alpha(k) = 1$ و $G = H^{-1}(x(k))$ (عکس ماتریس هسین تابع هزینه) در نظر بگیریم در این صورت رابطه نیوتن نتیجه خواهد شد. می‌بینیم که در این روش مشتق دوم بکار رفته است؛ بنابراین در نقاطی که تکینگی داشته باشد باعث مشکل خواهد شد.

روش گاوس-نیوتن^۵

در اغلب مسائل بهینه‌سازی، تابع هزینه، تابع حداقل مربعات خطا می‌باشد. به طور کلی می‌توان به صورت $I(x) = \sum_{i=1}^n f^2(i, x)$ آن را مطرح کرد. این نوع مسائل به حداقل مربعات غیرخطی^۶ معروف می‌باشند. برای این نوع توابع، روش گاوس-نیوتن که نوع خاص از روش نیوتن می‌باشد مناسب بوده و به صورت زیر بیان می‌شود:

$$X_k = X_{k-1} \alpha_{k-1} (J_{k-1}^T J_{k-1})^{-1} J_{k-1}^T f(x_{k-1}) \quad (7)$$

در آن J ماتریس ژاکوبین می‌باشد.

باید یادآور شویم که در عمل، محاسبه معکوس $J_{k-1}^T J_{k-1}$ صورت صریح نمی‌باشد و بجای آن از رابطه زیر می‌توان بردار جستجو را محاسبه نمود:

$$(J_{k-1}^T J_{k-1}) P_{k-1} = J_{k-1}^T f_{k-1} \quad (8)$$

البته ممکن است $J_{k-1}^T J_{k-1}$ مشکل تکنیکی داشته باشد و محاسبات را با مشکل مواجه نماید؛ لذا باید در این به طور مواقع از روش‌های دیگری مانند روش لوبنرک-مارکوآرت استفاده نمود.

روش گرادیان مزدوج

در روش‌های ذکر شده حافظه مورد نیاز به صورت کوادراتیک افزایش پیدا می‌کند و با افزایش تعداد پارامترها نیز پیچیدگی محاسبات بیشتر می‌شود. برای مسائل بزرگ اغلب تقریب زنی ماتریس هسین قدری مشکل می‌باشد. روش گرادیان مزدوج به صورت غیر مستقیم از تقریب ماتریس هسین جلوگیری می‌کند بنابراین حافظه مورد نیاز و پیچیدگی محاسبات در این روش به صورت خطی افزایش پیدا می‌کند.

تصادفی بین صفر و یک بر اساس یک توزیع یکنواخت تولید می‌شود. اگر این عدد از (احتمال برای عملگر "جهش-۲") کمتر باشد، مقدار پارامتر آن بیت که ممکن است وزن روی ارتباط‌های پیشخور یا پسخور یا بایاس باشد، صفر می‌شود. پس از تولید نسل جدید جستجویی به منظور تعیین بیت‌های غیر صفر هر عضو (کروموزوم) صورت می‌گیرد.

تابع هزینه اصلاح شده بکار رفته شده در روش الگوریتم ژنتیک اصلاح شده به صورت معادله ۳ می‌باشد. در این تابع معیار شایستگی هزینه‌ای بر تعداد وزن‌های غیر صفر اعمال می‌شود تا الگوریتم را در راستای رسیدن به ساختاری با حداقل ارتباط‌های ممکن تشویق نماید.

$$J = \sum_{i=1}^{n_d} (O_i - T_i)^2 + C \sqrt{\frac{\sum_{i=1}^{n_d} (O_i - T_i)^2}{n_d}} \quad (3)$$

$$Fit = \frac{1}{J} \quad (4)$$

در تابع فوق، C تعداد وزن‌های غیر صفر است که به صورت یک هزینه دینامیکی، یعنی متغیر با روند آموزش، اعمال شده است تا از تسلط آن بر معیار شایستگی جلوگیری شود. در صورتی که C را به صورت یک ضریب ثابت در معیار شایستگی قرار دهیم مشاهده می‌شود که در روند آموزش، مقدار MSE به مرور کاهش می‌یابد. در نتیجه، اهمیت خطای خروجی شبکه در معیار شایستگی از بین می‌رود و الگوریتم به سمت کوچک کردن شبکه (کاهش C) پیش رفته و آموزش نمی‌بیند. بنابراین، ضریب تأثیر تعداد وزن‌های غیر صفر باید متناسب با MSE کاهش یابد. در اصلاح شده، مجذور میانگین مربعات خطا (RMSE) به عنوان ضریب هزینه بر وزن‌های غیر صفر انتخاب شده است. این انتخاب به صورت آزمون و خطا بوده است و اثباتی برای بهینه‌گی آن وجود ندارد. بدین ترتیب هرچه مجذور میانگین مربعات خطا کاهش یابد، یعنی فرایند آموزش به مراحل آخر نزدیک شود، سهم هزینه بر وزن‌های غیر صفر ($C \times RMSE$) نیز کم‌رنگ تر می‌شود.

پس انتشار خطا

اغلب روش‌های آموزش شبکه‌های عصبی بر پایه الگوریتم‌های گرادینانی و نوعاً گونه‌ای از پس انتشار خطا برای یافتن پارامترهای وزنی شبکه می‌باشند. ساده‌ترین روش پیاده سازی الگوریتم آموزش پس انتشار خطا این است که وزن‌های و بایاس‌ها شبکه را در جهتی (منفی شیب) که در آن تابع هزینه یا عملکرد با سرعت بیشتری کاهش می‌یابد، به روز رسانی می‌شود. این نوع الگوریتم‌های آموزش موسوم به الگوریتم‌های گرادیان می‌باشد. روش‌های آموزش مبتنی بر گرادیان خطا، با استفاده از مشتق اول و یا دوم خطا، هر یک با به کارگیری رویه خاص خود، سعی می‌کنند با اصلاح پارامترهای سیستم، تابع معیار خطای تعریف شده برای آن را کمینه سازند. روش‌های بسیار متنوعه‌ای از الگوریتم پس انتشار وجود برای آموزش شبکه‌های عصبی وجود دارد. فرم کلی آن‌ها به صورت رابطه (۵) می‌باشد.

$$X_{next} = X_{present} + \alpha_{present} G.g_{present} = X(k) = X(k) + \alpha(k).G.g(x(k)) \quad (5)$$

در این رابطه X_k بردار وزن‌های شبکه و α_k بردار جستجو و نیز گام یا نرخ یادگیری می‌باشد.

این نوع روش‌ها به طور کلی شامل دو سری الگوریتم‌های نوع کند و الگوریتم‌های سریع می‌باشند. روش‌هایی از قبیل نزول گرادینانی، روش نیوتن روش گاوس-نیوتن از نوع روش‌های کند محسوب می‌شود و

³ - Levenberg - Marquardt

⁴ - Steepest Descend

⁵ - Gausse - Newton

⁶ -Nonlinear Least Square

آن جهت مخفی است که واحدهای آن تنها با نورون‌های درونی شبکه تعامل دارند و با دنیای خارج از شبکه مرتبط نیستند. در این شبکه لایه مخفی، شبکه را قادر می‌سازد برای هر ورودی اصلی شبکه، خروجی مطلوب تولید کند. نورون‌های لایه زمینه نیز دارای تاریخچه‌ای از حالت داخلی سیستم می‌باشند؛ لذا نورون‌های لایه پنهان علاوه بر ورودی اصلی شبکه، مجبور به نگاشت تاریخچه‌ای از حالت‌ها به خروجی هستند، در نتیجه شبکه اثر زمان را در فرایند درک می‌کند [۱]. ساختار این شبکه به عنوان یک شبکه چند لایه عمل می‌کند که لایه ورودی آن شامل ورودی‌های خارجی (ورودی‌های سیستم) و خروجی نورون‌های لایه زمینه می‌باشد.

شبکه المان با روابط زیر در این پروژه مورد استفاده قرار گرفته‌اند. در لایه میانی از تابع تبدیل تانژانت سیگموید و در لایه خروجی از تابع خطی استفاده شده است. نتایج استفاده از این نوع روابط در بهبود عملکرد شبکه، در بخش‌های بعد نشان داده خواهد شد.

$$Y_i^1(k) = \tan \operatorname{sig}(IW \times U + b^1 + CW^1 \times Y_i^1(k-1)) \quad (9)$$

$$Y_n^2(k) = \operatorname{purlin}(LW \times Y_i^1(k) + b^2) \quad (10)$$

$$U = [\delta_E \quad \delta_A \quad \delta_R]^T \quad (11)$$

شبکه المان اصلاح شده نیز همانند شبکه المان است (شکل ۱)، با این تفاوت که برای بالا بردن ظرفیت دینامیکی شبکه عناصر خود بازخورد در لایه زمینه به هر یک از نورون‌ها افزوده شده است. یعنی خروجی قبلی نورون‌های زمینه‌ای به خودش بازگشت داده می‌شود. در شبکه المان نورون‌های زمینه حالت‌های گذشته خروجی لایه میانی را نگه نمی‌دارند و فقط ارزش آخرین حالت ذخیره می‌شود. اما در شبکه المان اصلاح شده نورون‌های زمینه حالت‌های گذشته خروجی لایه میانی را حفظ می‌کنند؛ لذا حافظه دینامیکی شبکه افزایش می‌یابد. این نوع شبکه در شناسایی و شبیه‌سازی سیستم‌های دینامیکی هواپیما مناسب هستند [۱]. روابط استفاده شده برای این شبکه به صورت ذیل می‌باشد:

$$Y_i^1(k) = \tan \operatorname{sig} \left(IW \times U + b^1 + CW^1 \times Y_i^1(k-1) + \right. \\ \left. CW^2 \times Y_i^1(k-2) + CW^3 \times Y_i^1(k-3) \right) \quad (12)$$

$$Y_n^2(k) = \operatorname{purlin}(LW \times Y_i^1(k) + b^2) \quad (13)$$

استخراج داده پروازی

یک مرحله مهم در شناسایی سیستم استخراج داده‌های مناسب پروازی هواپیما است. داده‌های واقعی، نیز از داده‌های ثبت شده در آزمایشات پروازی یک هواپیما جنگنده نسل چهارم با قابلیت مانور بالا استفاده شده است. در این تحقیق از جنگنده مذکور با نام Xcraft یاد خواهد شد. آزمایش هواپیما در شرایط هوای آرام بدون باد و توربالانس با سرعت و ارتفاع معلوم صورت گرفته است. خلبان ورودی مناسب را به هر یک از ورودی‌های کنترلی هواپیما یعنی سه زاویه سطوح اصلی کنترلی (الویتور، ایلران و رادر) وارد می‌نماید. اعمال ورودی، در طول پرواز تست برای داده برداری، در شرایط تریم در یک ارتفاع و سرعت خاص توسط خلبان صورت گرفته است. البته معمولاً با توجه به قابلیت مانور بالا و چابکی هواپیمای مورد آزمایش اندکی تغییرات در سرعت و ارتفاع به هنگام آزمایش دیده می‌شود. در این نوع هواپیما، فرامین کنترلی دارای ارتباط داخلی با یکدیگر هستند؛ لذا زمانی که تنها یکی از فرامین به هواپیما وارد می‌شود، در دیگر سطوح کنترلی نیز اندکی تغییرات وجود خواهد داشت. داده‌های ثبت شده مورد استفاده برای شناسایی عبارتند از: شتاب طولی (n_x)، شتاب

در روش کوازی-نیوتن اگر بجای H_{k-1}^{-1} در هر تکرار مقدار ماتریس همانی I را قرار دهیم در این صورت ممکن است بردار جستجو P مستقیماً خود را تغییر دهد. از آنجا که روش گرادیان مزدوج از همان تقریب نامناسب روش کوازی-نیوتن استفاده می‌کند لذا معمولاً بردار جستجو ممکن است بدتر شود پس در این روش برای اینکه سریع‌تر همگرا شود نیاز به تکرار بیشتری خواهد داشت هرچند در کل مقدار زمان و هزینه محاسبه در روش گرادیان مزدوج برای مسائل بزرگ کمتر می‌باشد. الگوریتم روش گرادیان مزدوج به صورت زیر می‌باشد:

$$X_{k+1} = X_k + \alpha_k P_k \quad (6)$$

$$P_k = -g_k + B_k P_{k-1}$$

$$B_k = \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}}$$

مقدار اسکالر B اطلاعات تکرار قبلی را همراه خود دارد.

نوع دیگری از این روش توسط پولاک - رایبر y مطرح شد که بردار جستجو به صورت: $p_k - g_k + B_k P_{k-1}$ محاسبه می‌شود و مقدار اسکالر B_k نیز به صورت زیر می‌باشد:

$$B_k = \frac{\Delta g_{k-1}^T g_k}{g_{k-1}^T g_{k-1}} \quad (7)$$

روش لونبرگ-مارکوآرت

این روش شبیه کوازی-نیوتن می‌باشد و برای داشتن سرعت بیشتر بدون محاسبه ماتریس هسین طراحی شده است. در روش LM هم از مشتق اول (گرادیان) و هم مشتق دوم موسوم به هسین برای اصلاح پارامترها استفاده می‌شود. اگر تابع عملکرد به صورت مجموع مربعات خطا باشد، ماتریس هسین مقدار $H = J^T J$ و گرادیان نیز مقدار $g = J^T e$ خواهد شد. ماتریس ژاکوبین J نیز مشتق اول تابع عملکرد، و e بردار خطا در شبکه می‌باشد. پیچیدگی محاسبه ماتریس ژاکوبین از هسین کمتر می‌باشد.

الگوریتم لونبرگ - مارکوآرت به صورت زیر می‌باشد:

$$X_{k+1} = X_k - [J^T J + \alpha_k I]^{-1} J^T e \quad (8)$$

اگر α کوچک باشد الگوریتم تبدیل به الگوریتم گاوس - نیوتن خواهد شد. اگر α بزرگ باشد این روش به نزول تند تبدیل می‌شود.

روش نیوتن یک روش سریع و در نزدیک نقطه مینیمم دقت خوبی هم دارد؛ لذا در روش لونبرگ - مارکوآرت هدف این است که هر چه سریع‌تر به سمت روش نیوتن خود را سوق بدهد؛ بنابراین بهتر است در هر تکرار در صورتی که تابع عملکرد در حال کاهش می‌باشد، α نیز کاهش پیدا کند؛ لذا زمانی که α کاهش پیدا کند رفته رفته به سمت روش گاوس - نیوتن نزدیک می‌شویم که در این صورت در نزدیکی نقطه مینیمم دقت بالایی خواهد داشت.

شبکه عصبی در شناسایی سیستم‌های دینامیکی

ساختارهای مختلفی برای شناسایی دینامیک هواپیما در مراجع [۱ و ۲] معرفی شده است. با توجه این که هدف اصلی این مقاله معرفی الگوریتم آموزش مناسب می‌باشد تنها از شبکه المان اصلاح شده برای شناسایی استفاده شده است [۱]. شبکه المان یک شبکه چند لایه است که در سطح ورودی شبکه یک لایه مخفی به نام واحدهای زمینه وجود دارد. این لایه از

عرضی (n_y) و شتاب عمودی (n_z)، زوایای موقعیت پیچ (θ)، رول (ϕ) و زاویه یاو (ψ) زاویه حمله (AOA) و سرعت (v).

نتایج

به طور کلی هفت پارامتر $n_z, n_y, n_x, \psi, \theta, \phi, \alpha$ با یکی از روش‌های ذکر شده گرادیدانی و یا الگوریتم ژنتیک آموزش داده شد. شبکه استفاده شده نیز یک شبکه المان اصلاح شده می‌باشد که دارای ۸ نرون در لایه مخفی و دو تأخیر زمانی در لایه زمینه آن به صورت یک شبکه سه ورودی شامل ایلرون، الویتور و رادر و یک خروجی شامل هر یک از پارامترهای پروازی هواپیما، می‌باشد. یک مدل شناسایی شده در صورتی مدل مناسبی است که، به درستی قادر به پیش‌بینی رفتار دینامیکی سیستم برای ورودی‌های جدید باشد و یا به عبارت دیگر دارای اعتبارسنجی مناسبی برای ورودی‌های جدید داشته باشد و یا به عبارت دیگر دارای تعمیم‌دهی مناسبی نسبت به ورودی‌های جدید داشته باشد. روش‌های مختلفی برای انجام اعتبارسنجی مدل، به منظور استخراج یک مدل مناسب از اطلاعات جمع آوری شده، وجود دارد. ساده‌ترین روش برای تعیین اعتبار مدل، روش اعتبارسنجی متقابل^۸ است که در آن، داده‌ها به دو مجموعه جداگانه شامل، داده‌های آموزش و داده‌های اعتبارسنجی تقسیم شده‌اند. داده‌های آموزشی، به منظور تخمین پارامترهای مدل شناسایی شده، و داده‌های اعتبارسنجی، برای بررسی مدل، به منظور تعیین این نکته که آیا این مدل می‌تواند رفتار دینامیکی سیستم به خوبی پیش‌بینی کند، استفاده می‌شوند. شکل ۲ داده‌های تجربی آموزش و اعتبارسنجی را برای متغیر زاویه حمله با ورودی کنترلی الویتور نشان می‌دهد. این نمودار دارای دو بخش می‌باشد، یک بخش آن شامل داده‌هایی که برای آموزش مدل و به دست آورد ساختار مناسب مدل مورد استفاده قرار می‌گیرد و بخش دوم شامل داده‌هایی از تست پرواز است که برای اعتبارسنجی مدل به دست آمده مورد استفاده قرار می‌گیرد.

جدول‌های ۱ و ۴ میزان میانگین مربعات خطا (MSE) شبکه عصبی را برای متغیرهای مختلف پرواز برای داده‌های آموزش و تعمیم‌دهی با روش آموزش گرادیدانی و الگوریتم ژنتیک را نشان می‌دهد. آموزش برای هر یک از روش‌های با شرایط یکسان صورت گرفته است. بعضی روش‌های گرادیدانی از قبیل شیب تند قادر به آموزش شبکه در یادگیری دینامیک هواپیما نمی‌باشند. روند یادگیری در این روش‌ها کند بوده از طرفی در مینیمم‌های محلی نیز گرفتار می‌شوند بنابراین در اکثر موارد شبکه آموزش مناسبی نخواهد داشت. در این جداول نتایج روش‌های که قادر به آموزش شبکه می‌باشند، آورده شده است.

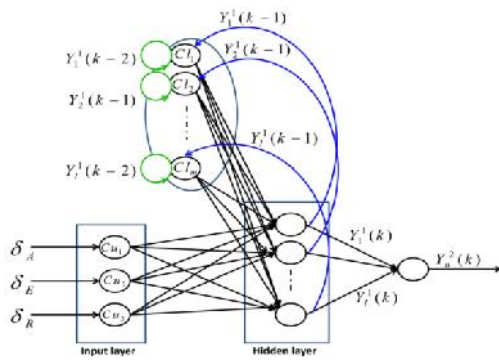
نکته‌ای که در مورد روش‌های گرادیدانی قابل اهمیت این است که در این روش‌ها با توجه به اینکه انتخاب اولیه پارامترها به صورت تصادفی می‌باشد در بعضی مواقع با تعداد اپک کم به مقدار هدف MSE مورد نظر می‌رسد. چنین مواقع‌های خطای تعمیم‌دهی خیلی کمتر از زمانی است که با تعداد اپک‌های بالاتر به MSE مورد نظر می‌رسد. برای مثال جدول ۳ تعداد اپک‌ها برای متغیرهای مختلف که خطای آن به انداز ۰/۰۲ کاهش یافته را نشان می‌دهد. جدول ۴ نتایج اعتباری سنجی هر یک از شبکه‌ها که به یک اندازه آموزش داده شده است را نشان می‌دهد. همان‌طور که مشخص است هر شبکه با تعداد اپک کمتری به میزان MSE مورد نظر (۰/۰۲) برسد خطای خروجی داده‌های تعمیم‌دهی کمتر می‌باشد. نتایج

نشان می‌دهد که روش لوبنرک روش سریع‌تری می‌باشد یعنی با تعداد اپک مساوی میزان MSE کمتری در روند آموزش کسب می‌کند، اما روش کوازی نیوتن در تعمیم‌دهی نتایج بهتری را دارا می‌باشد. یعنی خطای شبکه برای داده‌های جدید کمتر می‌باشد. حتی زمانی که میزان MSE برای داده‌های آموزش برای دو روش کوازی - نیوتن و لوبنرک یکسان باشد میزان خطای تعمیم‌دهی با روش کوازی - نیوتن کمتر می‌باشد؛ اما در مقایسه الگوریتم ژنتیک با روش‌های گرادیدانی، خطای شبکه‌ای که با الگوریتم ژنتیک آموزش دیده است کمتر می‌باشد. در الگوریتم ژنتیک مشکل تکینگی را نداریم فقط روند آموزش کندتر می‌باشد، به طوری که زمان آموزش بعضی از پارامترها چندین برابر روش‌های گرادیدانی می‌باشد. با توجه به این که در روش الگوریتم ژنتیک اصلاح شده، علاوه بر اینکه تابع هزینه بهینه می‌شود، ساختار شبکه نیز کوچک می‌شود و شبکه‌های کوچک‌تر دارای تعمیم‌دهی بهتری می‌باشند، زمانی که شبکه با این روش آموزش می‌بینید، الگوریتم ژنتیک اصلاح شده دارای خطای کمتری در مقایسه با سایر روش‌ها دارا می‌باشد؛ اما سرعت آموزش در این روش از سرعت یادگیری با الگوریتم ژنتیک معمولی نیز کمتر می‌باشد.

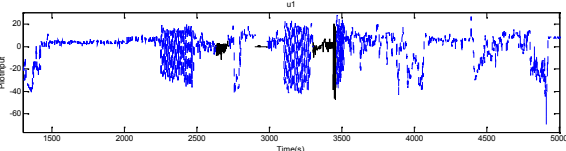
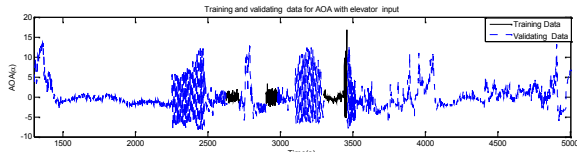
شکل ۳ توانایی الگوریتم ژنتیک در آموزش شبکه عصبی مورد استفاده نشان داده شده است که چگونه با گذشت نسل‌ها خروجی شبکه به خروجی مطلوب همگرا می‌شود. برای نمونه نتایج آموزش و تعمیم‌دهی بعضی از وضعیت‌های پروازی در شکل‌های ۵ الی ۸ آورده شده است. هر شکل دارای سه قسمت می‌باشد قسمت اول آن مقایسه مقدار هدف و مقدار خروجی از شبکه، قسمت دوم آن ورودی به شبکه و قسمت سوم هر شکل مقدار خطای خروجی شبکه و مقدار واقعی می‌باشد. نتایج این شکل-ها نشان می‌دهد روش‌های مذکور توانایی آموزش شبکه عصبی بازگشتی را در یادگیری رفتار دینامیکی هواپیما را دارد.

نتیجه‌گیری

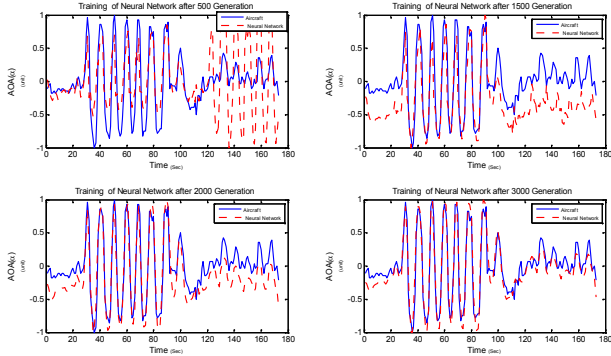
شبکه عصبی المان اصلاح شده به صورت یک شبکه سه ورودی و یک خروجی برای شناسایی و مدل‌سازی دینامیک غیرخطی هواپیما مورد استفاده قرار گرفت. آموزش شبکه با روش‌های مختلف آموزش پس انتشار خطا و الگوریتم ژنتیک صورت گرفت. نتایج نشان می‌دهد که بعضی از روش‌های گرادیدان از قبیل شیب تند روش مناسبی برای آموزش شبکه در شناسایی دینامیک هواپیما مناسب نمی‌باشد؛ اما روش‌های دیگر از قبیل لوبنرک و کوازی-نیوتن و گرادیدان مزدوج جز روش‌های یادگیری سریع می‌باشند که برای آموزش دینامیک هواپیما مناسب می‌باشند. در مقایسه روش‌های گرادیدانی با همدیگر مشخص گردید که سرعت آموزش شبکه‌های با لوبنرک آموزش می‌بیند بیشتر می‌باشد اما خطای شبکه تعمیم‌دهی روش کوازی نیوتن کمتر می‌باشد. الگوریتم ژنتیک نیز در مقایسه با روش‌های گرادیدانی سرعت بسیار کندی دارد اما زمانی که شبکه آموزش دید، خطای تعمیم‌دهی شبکه کمتر از روش‌های گرادیدانی می‌باشد. به طور کلی خطای تعمیم‌دهی شبکه‌ای که با الگوریتم ژنتیک اصلاح شده آموزش دیده کمتر از سایر شبکه‌ها در یادگیری رفتار دینامیکی هواپیما می‌باشد، اما سرعت یادگیری این روش نیز از سایر الگوریتم‌ها کمتر می‌باشد. بنابراین، با توجه به سرعت‌های آموزش شبکه در یادگیری رفتار هواپیما، می‌توان گفت الگوریتم ژنتیک برای آموزش آنلاین و الگوریتم لوبنرک و کوازی نیوتن علاوه بر یادگیری آفلاین برای یادگیری آنلاین رفتار دینامیکی هواپیما، برای کاربردهای مختلف مناسب می‌باشند.



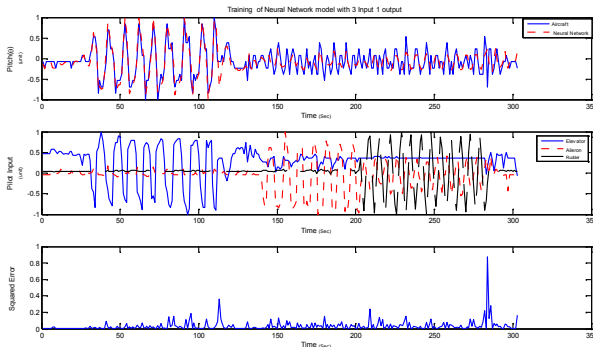
شکل ۱- معماری شبکه المان تغییر یافته



شکل ۲- داده‌های تست پرواز برای آموزش و اعتبارسنجی شبکه



شکل ۳- روند روبه بهبود شبکه با تولید نسل‌های مختلف و مقایسه خروجی شبکه با خروجی مطلوب



شکل ۵- نتیجه آموزش پارامتر پیچ با الگوریتم ژنتیک با سه ورودی

جدول ۱: مقایسه MSE نتایج آموزش شبکه عصبی با روش آموزش مختلف با یک‌های یکسان

پارامتر	LM	CG	QN
n_x	۰/۰۲۹۶	۰/۰۷۸۷	۰/۰۲۹۹
n_z	۰/۰۱۹۹	۰/۰۱۶۴	۰/۰۰۳۳
n_y	۰/۰۰۸۶	۰/۰۲۱۵	۰/۰۱۴۶
θ	۰/۰۱۱۹	۰/۰۵۹۹	۰/۰۲۲۸
φ	۰/۰۰۵۵	۰/۰۲۶۲	۰/۰۱۷۷
ψ	۰/۰۰۷۴	۰/۰۳۵۷	۰/۰۱۴۴
α	۰/۰۱۲۲	۰/۰۲۲۵	۰/۰۱۱۳

جدول ۲: مقایسه MSE نتایج تعمیم‌دهی شبکه عصبی با روش آموزش مختلف با یک‌های یکسان

پارامتر	LM	CG	QN
n_x	۰/۱۰۰۰	۰/۱۷۲۱	۰/۰۹۱۲
n_z	۰/۰۴۶۲	۰/۰۱۴۱	۰/۰۰۴۹
n_y	۰/۰۶۳۲	۰/۰۶۸۱	۰/۰۵۴۹
θ	۰/۰۷۸۰	۰/۱۴۹۷	۰/۰۶۶۱
φ	۰/۴۲۰۸	۰/۱۵۶۸	۰/۲۲۵۰
ψ	۰/۰۱۷۵	۰/۰۱۸۲	۰/۰۱۱۹
α	۰/۱۱۷۷	۰/۰۷۵۲	۰/۰۹۰۲

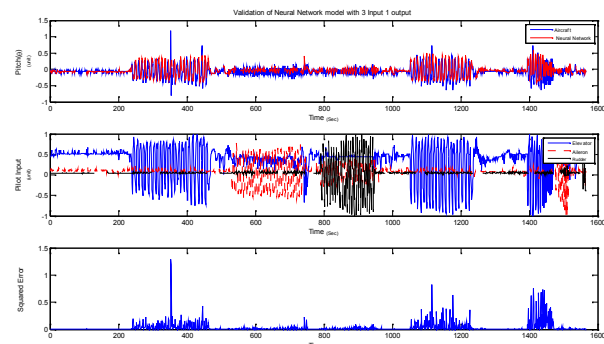
جدول ۳: تعداد یک‌های متغیرهای مختلف با MSE برای آموزش متغیرهای مختلف پرواز

پارامتر	LM	CG	QN
n_x	۴۳	۲۰۰	۳۲
n_z	۵	۱۲	۸
n_y	۵	۲۹	۵۷
θ	۱۰	۳۷۲	۵۱
φ	۱۰	۲۶۱	۱۸۷
ψ	۷	۵۷	۴۵
α	۱۰	۵۱	۲۹

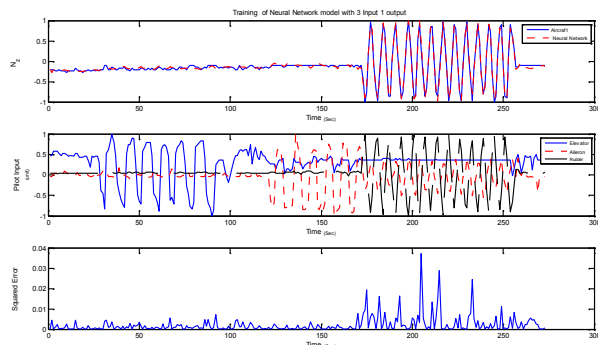
جدول ۴: مقایسه MSE نتایج تعمیم‌دهی شبکه عصبی با روش آموزش مختلف برای آموزش یکسان

پارامتر	NGA	MGA	LM	CG	QN
n_x	۰/۱۴۸۲	۰/۱۳۸۱	۰/۰۷۶۷	۰/۱۰۵۲	۰/۱۰۷۶
n_z	۰/۰۰۸۸	۰/۰۰۸۵	۰/۰۱۸۳	۰/۰۱۰۰	۰/۰۱۵۰
n_y	۰/۰۶۲۶	۰/۰۶۱۴	۰/۰۷۰۵	۰/۰۲۰۰	۰/۰۴۵۹
θ	۰/۰۳۷۳	۰/۰۳۵۱	۰/۰۹۵۱	۰/۱۰۴۵	۰/۰۷۷۶
φ	۰/۱۰۰۵	۰/۱۰۰۱	۰/۰۱۴۵	۰/۰۱۵۸	۰/۰۱۰۸
ψ	۰/۰۰۸۸	۰/۰۰۸۵	۰/۰۱۰۲	۰/۰۱۱۶	۰/۰۱۰۳
α	۰/۰۲۶۲	۰/۰۲۴۱	۰/۰۹۲۳	۰/۰۸۴۷	۰/۰۸۵۹

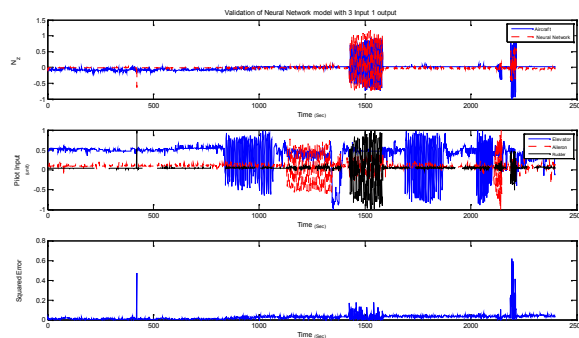
6. D. E. Rumelhart, G. E. Hinton and R. J. Williams, Learning representations by back-propagating errors, Nature, vol.323, no.6088, pp.533-536, 1986.
7. Sexton, R.S., Dorsey, R.E., Johnson, J.D., Toward Global Optimization of Neural Networks: A Comparison of the Genetic Algorithm and Backpropagation, Decision Support Systems 22, pp. 171-185, 1998.
8. Sexton R.S., Dorsey R.E., Sikander N.A., "Simultaneous Optimization of Neural Network Function and Architecture Algorithm", Decision Support Systems 36, 2004, PP.283-296.
9. Sexton, R.S., Dorsey, R.E., Sikander, N.A., Simultaneous Optimization of Neural Network Function and Architecture Algorithm, Decision Support Systems 36, pp.283-296, 2004.
10. Schaffer, J. D., Whitley, D., Eshelman, L. J., Combinations of genetic algorithms and neural networks: A survey of the state of the art, COGANN- 92. Combinations of Genetic Algorithms and Neural Networks. Los Alamitos, CA: IEEE Computer Society Press, pp. 1-37, 1992.



شکل ۶- نتیجه تعمیم‌دهی پارامتر پیچ با الگوریتم ژنتیک با سه ورودی



شکل ۷- نتیجه آموزش پارامتر N_z با سه ورودی با روش لونیگ



شکل ۸- نتیجه آموزش پارامتر N_z با سه ورودی با روش کوازی- نیوتن

مراجع

۱. ثقفی، فریبرز، و رودباری، علیرضا، "مدل‌سازی هوشمند دینامیک هواپیماهای جنگنده با استفاده از داده‌های ثبت شده در پرواز". ششمین کنفرانس هوافضای ایران ۵ تا ۷ اسفند ۱۳۸۵ تهران .
2. A. Roudbari and F. Saghafi, (2014). Intelligent Modeling and Identification of Aircraft Nonlinear Flight, doi: <http://dx.doi.org/10.1016/j.cja.2014.03.017>.
3. Saghafi F., Heravi, B.M., A New Approach to the Modeling of Highly Maneuver Aircrafts Using Neural Networks, Proc. 25th International Council of the Aeronautical Sciences (ICAS2006), Hamburg, Germany, 2006.
4. Bebis G., Georgiopoulos M., Kasparis T., Coupling Weight Elimination with Genetic Algorithms to Reduce Network Size and Preserve Generalization, Neurocomputing 17, pp. 167-194, 1997.
5. Castillo, P.A., Merelo, J.J., Prieto, A., Rivas, V., Romero, G., G-Prop: Global Optimization of Multilayer Perceptrons Using GAs, Neurocomputing 35, pp.149-163, 2000.

SID



ابزارهای
پژوهش



سرویس ترجمه
تخصصی



کارگاه های
آموزشی



بلاگ
مرکز اطلاعات علمی



سامانه ویراستاری
STES



فیلم های
آموزشی

کارگاه های آموزشی مرکز اطلاعات علمی



تازه های آموزش
آموزش مهارت های کاربردی در تدوین و چاپ مقالات ISI

آموزش مهارت های کاربردی
در تدوین و چاپ مقالات ISI



تازه های آموزش
روش تحقیق کمی

روش تحقیق کمی



تازه های آموزش
آموزش نرم افزار Word برای پژوهشگران

آموزش نرم افزار Word
برای پژوهشگران