

ارائه دو تابع جدید فعال ساز برای شبکه های عصبی مصنوعی با مقادیر مختلط

و کاربرد آن ها در مسائل با مقادیر حقیقی

ابوذر قربانی نژاد^۱

چکیده

انتخاب تابع فعال ساز مناسب یکی از چالش های اصلی شبکه های عصبی با مقادیر مختلط می باشد. تابع فعال ساز انتخابی، باید دو شرط مشتق پذیری و کران دار بودن را داشته باشد. زمانی که از این شبکه ها برای حل مسائل با مقادیر حقیقی استفاده می شود، تابع فعال ساز وظیفه نگاشت از فضای مختلط به فضای حقیقی را نیز بر عهده دارد. در این مقاله دو تابع جدید فعال ساز برای شبکه های با مقادیر مختلط پیشنهاد شده است که شروط فوق را به خوبی تأمین می کند. توابع فعال ساز پیشنهادی دارای چهار ناحیه اشباع بوده و بر خلاف پرسپترون دو لایه معمولی، توانایی حل مسائل جداناپذیر خطی را دارند. برای هر یک از توابع پیشنهادی، روابط اصلاح وزن استخراج شده و روند آموزش و تست شبکه عصبی شرح داده شده است. با استفاده از دو مجموعه داده ی تشخیص پزشکی، مربوط به بیماری های دیابت و سرطان سینه، عملکرد شبکه بر روی مسائل با مقادیر حقیقی مورد سنجش قرار گرفته و نشان داده شده است که شبکه عصبی مختلط با توابع پیشنهادی، ساختار ساده تر و سرعت همگرایی بالاتری نسبت به شبکه های چند لایه پرسپترون استاندارد دارند. میزان صحت تشخیص این شبکه ها، برای بیماری دیابت ۸۰٪ و برای سرطان سینه ۹۵٪ می باشد.

کلیدواژه

تابع فعال ساز، تشخیص بیماری، دیابت، سرطان سینه، شبکه عصبی با مقادیر مختلط.

^۱ کارشناس ارشد مهندسی برق الکترونیک، دانشگاه تربیت مدرس aboo.gh@gmail.com

تاریخ پذیرش: ۱۵ تیر ۱۳۹۱

تاریخ دریافت: ۱۷ فروردین ۱۳۹۱

اهمیت بالایی که اعداد مختلط در زمینه‌های مختلف علوم دارند، سبب شد تا نوع جدیدی از شبکه‌های عصبی مصنوعی که بتوانند مستقیماً بر روی مسائل با مقادیر مختلط عمل نمایند پیشنهاد شوند [۱-۳]. با توجه به اینکه وزن‌ها و مقادیر ورودی و خروجی در این نوع شبکه‌ها مقادیر مختلط می‌باشند، این شبکه‌ها، شبکه‌های عصبی با مقادیر مختلط نامیده شده‌اند. این شبکه‌ها از دهه ۱۹۸۰ به بعد بسیار مورد توجه قرار گرفته و در هر دو حوزه تئوری و کاربردی توسعه یافتند، به طوری که امروزه کاربردهای بسیاری در زمینه‌هایی همچون: پردازش تصاویر دیجیتال و تصاویر راداری، الکترونیک نوری و پردازش اطلاعات پیدا کرده‌اند [۴-۶]. استفاده از وزن‌های مختلط در این شبکه‌ها، ویژگی‌های مطلوب زیر را نسبت به شبکه‌های عصبی با مقادیر حقیقی، به وجود آورده است:

- ۱- سرعت متوسط همگرایی در شبکه‌های عصبی با مقادیر مختلط، دو یا سه برابر شبکه‌های عصبی با مقادیر حقیقی است.
- ۲- تعداد لایه‌های مخفی مورد نیاز در این نوع از شبکه‌ها تقریباً نصف تعداد لایه‌های مخفی مورد نیاز در شبکه‌های با مقادیر حقیقی است [۳].

در نوروهای شبکه‌های عصبی مصنوعی، معمولاً حاصل جمع ورودی‌ها، از یک تابع فعال‌ساز عبور کرده و خروجی متناسب با مجموع ورودی‌ها برای نرون مشخص می‌گردد. در شبکه‌های عصبی با مقادیر حقیقی، انتخاب یک تابع فعال‌ساز با دو معیار پیوستگی و کران دار بودن صورت می‌گیرد که در مرجع [۷] انواع مختلف توابع فعال‌ساز برای این شبکه‌ها معرفی و از نظر عملکرد با یکدیگر مقایسه شده‌اند. اما در حوزه اعداد مختلط بر اساس تئوری لیویل^۲ [۸] هیچ تابع مختلطی که بتواند به طور هم‌زمان در کل صفحه مختلط، مشتق‌پذیر و کران دار باشد، وجود ندارد مگر اینکه برابر با عدد ثابتی باشد. به این دلیل انتخاب یک تابع فعال‌ساز مناسب، یکی از چالش‌های مهم برای طراحی شبکه‌های عصبی با مقادیر مختلط می‌باشد [۹]. تاکنون محققان توابع فعال‌ساز مختلفی را برای این نوع از شبکه‌ها ارائه نموده‌اند و برای اینکه بتوانند مسأله مشتق‌پذیری و کران دار بودن را حل نمایند، به ناچار محدودیت‌هایی برای داده‌های ورودی قرار داده‌اند. به عنوان مثال مرجع [۱۰] از تابع سیگموید به صورت $f(z) = 1/(1 - \exp(-z))$ استفاده نموده، اما با توجه به اینکه این تابع

^۲ Liouville

در نقاط $z = (2n+1)i\pi, n \in \mathbb{Z}$ دارای نقطه‌ی تکین می‌باشد، طراح مجبور شده است تا ورودی‌ها را در فضای خاصی از صفحه مختلط محدود نماید. توابع فعال‌ساز دیگری نیز پیشنهاد شده‌اند که هر یک به گونه‌ای محدودیت‌هایی بر روی ورودی‌ها داشته‌اند. این توابع به صورت کامل در مرجع [۵] بررسی و با یکدیگر مقایسه شده‌اند.

یک پرسپترون عادی دو لایه توانایی حل مسائلی که به صورت خطی جدا پذیر باشند را ندارد و به همین دلیل ساختارهای چند لایه پرسپترون پیشنهاد شدند. انتخاب مناسب تعداد لایه‌های مخفی و همچنین تعداد نورون‌های هر لایه مخفی یکی از مسائلی است که عملکرد شبکه را از نظر زمان آموزش، حداقل خطای شبکه، توانایی یادگیری و دیگر عوامل تحت تأثیر قرار می‌دهد. معمولاً معیاری ساده و صریح برای این انتخاب وجود ندارد و با سعی و خطا می‌توان به نتیجه مطلوب دست یافت. اما اخیراً دو محقق، شبکه‌های عصبی با مقادیر مختلط را بر روی مسائل دسته بندی با مقادیر غیر مختلط (حقیقی) به کار برده‌اند [۱۲، ۱۱]. همان‌گونه که در این مقاله نشان خواهیم داد، استفاده از وزن‌های مختلط در این شبکه‌ها سبب به وجود آمدن فضایی دو بعدی (فضای حقیقی و موهومی) برای شبکه شده و شبکه (بر خلاف یک پرسپترون معمولی) توانایی حل مسائلی که به صورت خطی جدا پذیر نیستند را دارا می‌باشد. این خاصیت مهم، سبب می‌شود تا با استفاده از یک ساختار ساده دو لایه، مسائل دسته بندی واقعی را حل نماید و مشکل انتخاب تعداد لایه‌های پنهان و مسائل مربوط به آن بر طرف گردد.

توابع فعال‌ساز شبکه‌های عصبی با مقادیر مختلط، زمانی که شبکه برای حل مسائل با مقادیر حقیقی استفاده می‌شود، علاوه بر ویژگی‌های مذکور، تبدیل مناسب فضای مختلط (مجموع وزن دار ورودی‌های مختلط) به فضای حقیقی را نیز بر عهده دارند. با توجه به تابع فعال‌ساز ارائه شده در [۱۱]، الگوریتم یادگیری شبکه شامل عبارات معکوس مشتقات جزئی می‌باشد. لذا فرایند یادگیری در زمانی که مشتقات جزئی صفر شوند، ناپایدار می‌گردد. در [۱۲] دو تابع جدید فعال‌ساز معرفی شده‌اند که مشکلات مربوط به توابع فعال‌ساز را تا حدودی برطرف نموده‌اند. توابع فعال‌ساز مطرح شده این مرجع، دارای چهار ناحیه اشباع می‌باشند، که در تابع فعال‌ساز اول، یکی از این نواحی غیر قابل استفاده می‌باشد که سبب کاهش عملکرد شبکه در دسته‌بندی ورودی می‌باشد.

در ادامه به ارائه دو تابع جدید فعال‌ساز خواهیم پرداخت که قسمت‌های موهومی و حقیقی را به صورت مناسبی با یکدیگر ترکیب نموده و از هر چهار ناحیه اشباع استفاده می‌نماید و نشان خواهیم داد که شبکه‌های مصنوعی با مقادیر مختلط، که از توابع فعال‌ساز پیشنهاد شده استفاده می‌نمایند،

سرعت همگرایی بیشتر و خطای دسته بندی کمتری نسبت به شبکه های عصبی متناظر با مقادیر حقیقی خواهند داشت. بدین منظور، از دو مجموع داده مربوط به بیماری دیابت و سرطان که از نوع مسائل دسته بندی می باشند استفاده می نماییم.

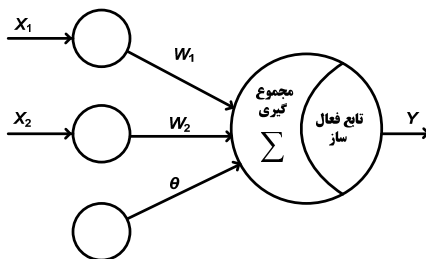
در ادامه این مقاله، ابتدا به معرفی شبکه عصبی پیشنهادی پرداخته و نحوه تبدیل ورودی های حقیقی به مقادیر مختلط را توضیح می دهیم. سپس توابع فعال ساز پیشنهادی و روابط اصلاح وزن ها در مرحله یادگیری شبکه را ارائه می کنیم. در ادامه، به نحوه آموزش شبکه پرداخته و سرانجام، نتایج شبیه سازی بدست آمده از شبکه در تشخیص بیماری های دیابت و سرطان سینه را ارائه می کنیم.

معرفی شبکه عصبی بکار رفته

همان طوری که بیان شد، شبکه عصبی بکاررفته در این مقاله، دارای وزن های مختلط است. با توجه به اینکه داده های ورودی این شبکه، حقیقی هستند، برای بهره گیری از ویژگی های بیان شده، باید قبل از اعمال ورودی ها به شبکه، آن ها را به یک عدد مختلط تبدیل کرد. در نوروں های لایه خروجی نیاز به تابعی است که عدد مختلط (مجموع حاصل ضرب وزن های مختلط در مقادیر نوروں های لایه ورودی) را به فضای حقیقی نگاشت کند. به این تابع، به اصطلاح تابع فعال ساز گفته می شود. در این مقاله دو تابع فعال ساز پیشنهاد می کنیم که فضای مختلط را به فضای حقیقی نگاشت می کنند. وظیفه این توابع، تقسیم فضای ورودی (مجموع وزن دار ورودی ها) به چند ناحیه برای نشان دادن کلاس های مختلف است. چون فضای ورودی به این توابع، عددی مختلط است، پس فضای ورودی یک فضای دو بعدی است (حقیقی و موهومی). با توجه به اینکه توابع فعال ساز مذکور نسبت به قسمت حقیقی و موهومی ورودی (مجموع وزن دار ورودی ها) مشتق پذیرند؛ می توان از الگوریتم گرادیان پایه^۳ روابط اصلاح وزن ها را بدست آورد.

در شکل ۱ ساختار یک شبکه عصبی ساده با مقادیر مختلط نشان داده شده است. همان گونه که مشاهده می شود این شبکه بسیار شبیه پرسپترون دو لایه است، با این تفاوت که وزن های شبکه به صورت مختلط است. در این شکل مشاهده می شود که ورودی ها و نوروں بایاس در وزن های مختلط متناظر ضرب و سپس مجموع آن ها از تابع فعال ساز عبور کرده و خروجی با مقدار حقیقی متناسب با ورودی، تولید می گردد.

^۳ gradient-based



شکل ۱. مدل ساده‌ای از شبکه‌های با مقادیر مختلط. تابع فعال‌ساز پس از محاسبه مجموع حاصل ضرب وزن‌ها در مقادیر ورودی، خروجی را به صورت عدد حقیقی متناسب با کلاس ورودی تعیین می‌کند.

تبدیل ورودی‌های حقیقی به اعداد مختلط

یکی از خواص اعداد حقیقی این است که می‌توان به راحتی در مورد بزرگ‌تر یا کوچک‌تر بودن آن‌ها نسبت به یکدیگر بحث کرد. برای نمونه در مورد دو عدد حقیقی a و b ، عدد a بزرگ‌تر، مساوی، یا کوچک‌تر از عدد b است. اما در مورد اعداد مختلط چنین نیست. به عنوان مثال برای دو عدد مختلط c و d عبارت $c > d$ صحیح نیست. اما می‌توان در مورد اندازه و فاز اعداد مختلط این تعریف را به کار برد؛ یعنی می‌توان گفت که $|c| > |d|$ یا $\text{phase}(c) > \text{phase}(d)$. اگر بخواهیم ورودی‌های شبکه را به عدد مختلط تبدیل کنیم، به طوری که خاصیت بزرگی و یا کوچکی ورودی‌ها نیز رعایت شود، یکی از راه‌ها این است که فاز را ثابت نگه داشته و اندازه آن‌ها را به تناسب تغییر دهیم. راه دیگری که در این مقاله از آن استفاده می‌کنیم، این است که اندازه را ثابت گرفته و فاز را متناسب با ورودی تغییر می‌دهیم. به طوری که همه ورودی‌ها به صورت اعداد مختلط با اندازه واحد و فازهای متفاوت کد شوند. با این فرض که ورودی‌های حقیقی در بازه $[a, b]$ ، توسط رابطه (۱) به صورت فازهایی بین $[0, \pi]$ تبدیل می‌شوند.

$$\phi = \frac{\pi(x-a)}{(b-a)} \quad (1)$$

اکنون فاز ϕ بدست آمده را در رابطه اولر (۲) قرار داده تا ورودی حقیقی x به عدد مختلط z تبدیل شود.

$$z = e^{i\phi} = \cos \phi + i \sin \phi \quad (2)$$

این تبدیل دارای مزایای مناسبی می‌باشد که می‌توان به نکات زیر در مورد آن اشاره کرد:

- این تبدیل خاصیت بزرگی و کوچکی بین ورودی‌ها را حفظ می‌کند برای مثال اگر دو عدد حقیقی x_1 و x_2 دارای رابطه $x_1 > x_2$ باشند اعداد مختلط نظیر آن‌ها نیز دارای همین رابطه به صورت $\text{phase}(z_1) > \text{phase}(z_2)$ می‌باشند.
- فاصله مکانی میان ورودی‌ها نیز حفظ شده است. برای مثال اگر دو عدد x_1 و x_2 هنگامی که $x_1 = a$ و $x_2 = b$ بیشترین فاصله را از یکدیگر داشته باشند، اعداد مختلط نظیر آن‌ها نیز بیشترین فاصله را از یکدیگر خواهند داشت.
- فاصله $[0, \pi]$ بهتر از فاصله $[0, 2\pi]$ می‌باشد زیرا در فاصله $[0, 2\pi]$ فاصله مکانی بین ورودی‌ها حفظ نخواهد شد و کوچک‌ترین و بزرگ‌ترین اعداد به یک نقطه نگاشت می‌شوند.

توابع فعال‌ساز

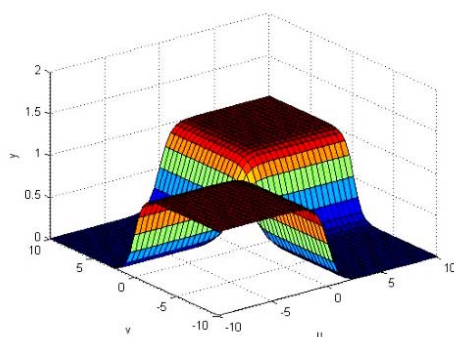
معمولاً توابع فعال‌ساز شبکه‌های با مقادیر حقیقی، در تمام بازه اعداد حقیقی مشتق‌پذیر و کران‌دار هستند، در نتیجه می‌توان الگوریتم گرادیان پایه را برای اصلاح وزن‌های شبکه، در آن‌ها به کار برد. اما طبق تئوری لیویل هیچ تابع مختلطی وجود ندارد که به طور هم‌زمان هم مشتق‌پذیر و هم کران‌دار باشد؛ مگر اینکه برابر با یک مقدار ثابت باشد [۸]. برای اینکه بتوان از الگوریتم گرادیان پایه در شبکه عصبی مختلط بهره برد، می‌توان قسمت‌های حقیقی و موهومی خروجی‌های مختلط را به صورت دو تابع حقیقی مجزا در نظر گرفت و سپس با استفاده از تابعی تحت عنوان تابع فعال‌ساز، آن‌ها را با یکدیگر ترکیب نمود، به صورتی که تابع بدست آمده هم مشتق‌پذیر بوده و هم کران‌دار باشد. در این بخش دو تابع جدید فعال‌ساز معرفی می‌شوند که بخش‌های حقیقی و موهومی (مربوط به مجموع وزن دار ورودی‌ها) را به صورت مناسبی ترکیب نموده و علاوه بر مشتق‌پذیر و کران دار بودن، وظیفه نگاشت از فضای مختلط را به فضای حقیقی به خوبی انجام خواهند داد. دو تابع فعال‌ساز پیشنهادی که وظیفه نگاشت از فضای موهومی به حقیقی را بر عهده دارند، در روابط (۳) و (۴) معرفی شده‌اند که در آن‌ها، مجموع حاصل ضرب ورودی‌ها در وزن‌های مختلط، در یک نورون به صورت $z = u + iv$ در نظر گرفته شده است.

$$y = f_{C \rightarrow R}(z) = \frac{1}{2}(1 + f_R(u) \times f_R(v)) \quad (۳)$$

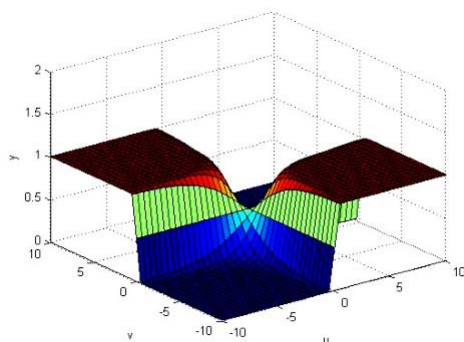
$$y = f_{C \rightarrow R}(z) = \frac{1}{1 + e^{uv}} \quad (۴)$$

که در رابطه (۳)، $f_R(x) = \tanh(x)$ است.

شکل‌های ۲ و ۳ به ترتیب، مقادیر خروجی متناظر با توابع فعال‌ساز (۳) و (۴) را بر حسب مقادیر مختلف u و v (قسمت‌های حقیقی و موهومی Z) نشان می‌دهند. تابع سیگموید در شبکه‌های عصبی معمولی دارای دو ناحیه اشباع می‌باشد، اما همان‌گونه که در شکل‌های ۲ و ۳ مشاهده می‌شود، هر دو تابع فعال‌ساز پیشنهاد شده دارای چهار ناحیه اشباع می‌باشند. در ادامه نشان داده خواهد شد که وجود این چهار ناحیه اشباع سبب می‌شود تا قابلیت این شبکه‌ها به طور قابل توجهی افزایش یابد.



شکل ۲. خروجی تابع فعال‌ساز پیشنهادی اول (رابطه (۳)) بر حسب مقادیر مختلف u و v .



شکل ۳. خروجی تابع فعال‌ساز پیشنهادی دوم (رابطه (۴)) بر حسب مقادیر مختلف u و v .

رابطه اصلاح وزن‌ها (یادگیری شبکه عصبی)

در این قسمت، رابطه اصلاح وزن‌ها را با استفاده از الگوریتم گرادیان پایه بیان می‌کنیم. با فرض یک شبکه عصبی مختلط با ابعاد $m \times n$ که m تعداد نورون‌های لایه ورودی و n تعداد نورون‌های لایه خروجی است. همچنین فرض می‌کنیم که x_j مقدار نورون j ام ورودی، w_{kj} وزن بین نورون k ام خروجی با نورون j ام ورودی و θ_k وزن بین نورون بایاس و نورون k ام خروجی باشند. برای نشان دادن قسمت‌های حقیقی و موهومی هر کدام از پارامترهای بالا آن‌ها را به صورت $x_j = x_j^R + ix_j^I$ و $\theta_k = \theta_k^R + i\theta_k^I$ و $w_{kj} = w_{kj}^R + iw_{kj}^I$ حاصل ضرب وزن‌ها در ورودی‌ها برای نورون k ام بوده و y_k مقدار خروجی نورون k ام پس از عبور از تابع فعال‌ساز باشد، داریم:

$$z_k = \sum_{j=1}^m w_{kj} x_j + \theta_k \quad (5)$$

$$z_k = \left(\sum_{j=1}^m (w_{kj}^R x_j^R - w_{kj}^I x_j^I) + \theta_k^R \right) + i \left(\sum_{j=1}^m (w_{kj}^I x_j^R + w_{kj}^R x_j^I) + \theta_k^I \right) \quad (6)$$

$$z_k = z_k^R + iz_k^I \quad (7)$$

$$y_k = f_{C \rightarrow R}(z_k) \quad (8)$$

اگر مقدار مطلوب نورون k ام، برابر d_k باشد تابع خطایی که باید در طول آموزش مینیمم شود عبارت است:

$$E = \left(\frac{1}{2} \right) \sum_{k=1}^n e_k^2 = \left(\frac{1}{2} \right) \sum_{k=1}^n (d_k - y_k)^2 \quad (9)$$

وزن‌ها و بایاس‌های شبکه توسط روابط (۱۰) تا (۱۵) در طول آموزش به روزرسانی می‌شوند.

$$\Delta W_{kj} = -\eta \left(\frac{\partial E}{\partial w_{kj}^R} + i \frac{\partial E}{\partial w_{kj}^I} \right) = \Delta w_{kj}^R + i \Delta w_{kj}^I \quad (10)$$

$$\Delta w_{kj}^R = -\eta \frac{\partial E}{\partial w_{kj}^R} = -\eta \frac{\partial E}{\partial e_k} \cdot \frac{\partial e_k}{\partial y_k} \cdot \left[\left(\frac{\partial y_k}{\partial f_R(z_k^R)} \cdot \frac{\partial f_R(z_k^R)}{\partial z_k^R} \cdot \frac{\partial z_k^R}{\partial w_{kj}^R} \right) + \left(\frac{\partial y_k}{\partial f_R(z_k^I)} \cdot \frac{\partial f_R(z_k^I)}{\partial z_k^I} \cdot \frac{\partial z_k^I}{\partial w_{kj}^R} \right) \right] \quad (11)$$

$$\Delta w_{kj}^I = -\eta \frac{\partial E}{\partial w_{kj}^I} = -\eta \frac{\partial E}{\partial e_k} \cdot \frac{\partial e_k}{\partial y_k} \cdot \left[\left(\frac{\partial y_k}{\partial f_R(z_k^R)} \cdot \frac{\partial f_R(z_k^R)}{\partial z_k^R} \cdot \frac{\partial z_k^R}{\partial w_{kj}^I} \right) + \left(\frac{\partial y_k}{\partial f_R(z_k^I)} \cdot \frac{\partial f_R(z_k^I)}{\partial z_k^I} \cdot \frac{\partial z_k^I}{\partial w_{kj}^I} \right) \right] \quad (12)$$

$$\Delta \theta_k = -\eta \left(\frac{\partial E}{\partial \theta_k^R} + i \frac{\partial E}{\partial \theta_k^I} \right) = \Delta \theta_k^R + i \Delta \theta_k^I \quad (13)$$

$$\Delta \theta_k^R = -\eta \frac{\partial E}{\partial \theta_k^R} = -\eta \frac{\partial E}{\partial e_k} \cdot \frac{\partial e_k}{\partial y_k} \cdot \left[\left(\frac{\partial y_k}{\partial f_R(z_k^R)} \cdot \frac{\partial f_R(z_k^R)}{\partial z_k^R} \cdot \frac{\partial z_k^R}{\partial \theta_k^R} \right) + \left(\frac{\partial y_k}{\partial f_R(z_k^I)} \cdot \frac{\partial f_R(z_k^I)}{\partial z_k^I} \cdot \frac{\partial z_k^I}{\partial \theta_k^R} \right) \right] \quad (14)$$

$$\Delta \theta_k^I = -\eta \frac{\partial E}{\partial \theta_k^I} = -\eta \frac{\partial E}{\partial e_k} \cdot \frac{\partial e_k}{\partial y_k} \cdot \left[\left(\frac{\partial y_k}{\partial f_R(z_k^R)} \cdot \frac{\partial f_R(z_k^R)}{\partial z_k^R} \cdot \frac{\partial z_k^R}{\partial \theta_k^I} \right) + \left(\frac{\partial y_k}{\partial f_R(z_k^I)} \cdot \frac{\partial f_R(z_k^I)}{\partial z_k^I} \cdot \frac{\partial z_k^I}{\partial \theta_k^I} \right) \right] \quad (15)$$

در روابط بالا η نرخ یادگیری سیستم است. با توجه به محاسبات، مقدار Δw_{kj} برابر است با [۱۲]:

$$\Delta w_{kj} = \bar{x}_j \Delta \theta_k \quad (16)$$

که $\bar{x}_j$ در رابطه بالا برابر با مزدوج مختلط مقدار x_j است.

برای تابع فعال ساز اول (رابطه (۳)) داریم:

$$\Delta \theta_k^I = \eta e_k f_R'(z_k^R) f_R'(z_k^I) \quad (17)$$

که در آن $f_R(u) = \frac{1}{2} \tanh(u)$ و $f_R'(u) = \frac{1}{2 \cosh^2(u)}$ است.

در تابع فعال ساز دوم (رابطه (۴)) نیز، روند اصلاح وزن نورون های بایاس توسط روابط (۱۸) و (۱۹) بیان می شود.

$$\Delta \theta_k^R = \eta e_k f_R'(z_k^R) \quad (18)$$

$$\Delta \theta_k^I = \eta e_k f_R'(z_k^I) \quad (19)$$

در روابط بالا، مقدار $f_R'(x)$ از مشتق گیری مستقیم از رابطه (۴) محاسبه می شود. برای تمام توابع فعال ساز مقدار Δw_{kj} با استفاده از رابطه (۱۶) محاسبه شده و اصلاح وزن انجام می شود.

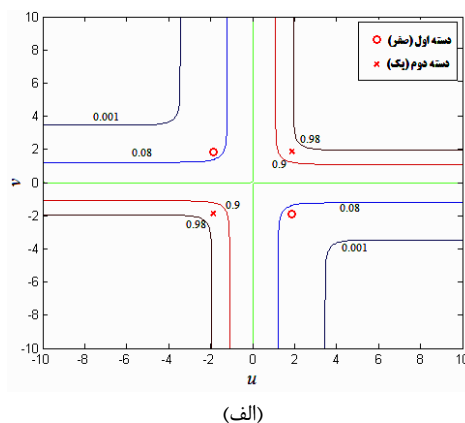
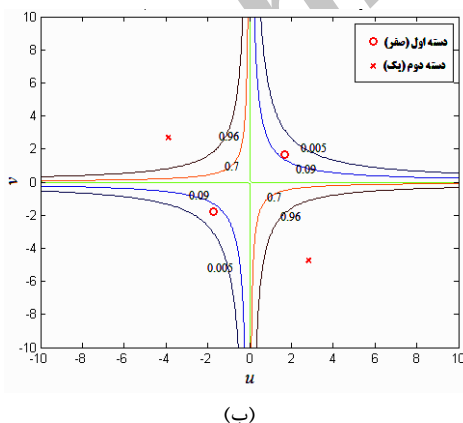
قابلیت حل مسائل جداناپذیر خطی

عدم توانایی حل مسائل جداناپذیر خطی، یکی از مشکلات اساسی یک پرسپترون عادی دو لایه است. برای حل این مشکل اساسی، ساختارهای چند لایه پرسپترون پیشنهاد شدند. انتخاب مناسب تعداد لایه های مخفی و همچنین تعداد نورون های هر لایه مخفی یکی از مسائلی است که عملکرد شبکه را از نظر زمان آموزش، حداقل خطای شبکه، توانایی یادگیری و دیگر عوامل تحت تأثیر قرار می دهد. در این بخش عملکرد شبکه های عصبی با مقادیر مختلط، که از توابع فعال ساز پیشنهاد شده بهره برده اند را نمایش دهیم. بدین منظور توسط این شبکه ها، مسأله XOR را که جزء مسائل جداناپذیر خطی می باشد، حل می نماییم. جدول ۱ ورودی های قبل و بعد از تبدیل و خروجی های مطلوب، که شبکه باید آموزش ببیند، آورده شده اند.

جدول ۱. مقادیر ورودی و خروجی تابع جدا ناپذیر خطی XOR.

خروجی	ورودی‌ها بعد از تبدیل		ورودی‌ها قبل از تبدیل	
	X_2	X_1	B	A
0	$1+i0$	$1+i0$	0	0
1	$-1+i0$	$1+i0$	1	0
1	$1+i0$	$-1+i0$	0	1
0	$-1+i0$	$-1+i0$	1	1

پس از آموزش شبکه توسط روابطی که در بخش قبل نشان داده شد، مقدار $z = w_1x_1 + w_2x_2 + \theta = u + iv$ مجموع حاصل ضرب ورودی‌ها در وزن‌های مختلط را نشان می‌دهد که پس از عبور از هر یک از توابع فعال‌ساز پیشنهاد شده، خروجی نهایی شبکه بدست می‌آید. شکل ۴-الف و ۴-ب به ترتیب نمودارهای کانتور توابع فعال‌ساز اول (رابطه (۳)) و دوم (رابطه (۴)) را نشان می‌دهند. در این شکل‌ها همچنین مجموع وزن‌دار ورودی‌ها (یعنی Z) برای هر چهار ورودی مختلف نیز نمایش داده شده است. برای تابع فعال‌ساز اول در شکل ۴-الف مشاهده می‌شود که شبکه پس از آموزش، مقادیری که در ربع اول و سوم قرار گرفته‌اند را به مقدار یک و خروجی‌هایی که در نواحی ربع دوم و چهارم قرار گرفته‌اند را به صفر نگاشت می‌کند. در شکل ۴-ب نیز این مسئله قابل مشاهده است با این تفاوت که خروجی‌هایی که در ربع دوم و چهارم قرار گرفته‌اند به یک و خروجی‌های دو ربع دیگر، به صفر نگاشت شده‌اند. لذا دیده می‌شود که این شبکه‌ها قادر به حل مسئله XOR که جزء مسائل جداناپذیر خطی است، می‌باشند.



شکل ۴. نمودارهای کانتور توابع فعال‌ساز اول و دوم و خروجی‌های مربوط به مسئله جداناپذیر خطی XOR.

روش آموزش شبکه عصبی پیشنهادی

روند اعمال داده‌ها به شبکه عصبی مذکور بدین صورت است که ابتدا همه داده‌ها به سه قسمت: داده‌های آموزشی^۴، داده‌های تأییدی^۵ و داده‌های آزمونی^۶ تقسیم می‌شوند. طریقه اعمال داده‌ها به شبکه به این صورت است که پس از هر پنج مرتبه آموزش شبکه توسط داده‌های آموزشی، یک بار داده‌های تأییدی به شبکه اعمال می‌شوند (بدون اصلاح وزن‌ها) و مقدار متوسط خطا توسط رابطه (۲۰) محاسبه شده و مقادیر وزن‌ها و بایاس‌ها در این مرحله ذخیره می‌گردند.

$$E = \frac{1}{2P} \sum_{p=1}^P \sum_{i=1}^n (d_{pi} - y_{pi})^2 \quad (20)$$

که p شماره نمونه، P تعداد کل نمونه‌ها و n تعداد نورون‌های لایه خروجی است. سپس از بین متوسط خطاهای محاسبه شده در مرحله‌ی تأیید، مقدار کمینه انتخاب شده و وزن‌ها و بایاس‌های متناظر، به عنوان وزن‌ها و بایاس‌های نهایی برای شبکه تثبیت می‌گردند. بعد از تعیین وزن‌های شبکه، داده‌های آزمونی به شبکه اعمال شده و میزان خطای دسته بندی، توسط شبکه، اندازه‌گیری می‌شود و درصد موفقیت شبکه در تشخیص صحیح نمونه‌های آموزش دیده نشده را با استفاده از رابطه (۲۱) محاسبه می‌نماییم.

$$(21) \quad \left[\text{تعداد کل داده‌های آزمونی} / (\text{تعداد کل داده‌های آزمونی} - \text{میزان خطای دسته بندی}) \right] \times 100 = \text{درصد موفقیت}$$

نتایج عملی

برای آزمون عملکرد هر دو شبکه پیشنهادی، از دو مجموعه داده در PROBEN1 استفاده شد [۱۳]، که در ادامه به طور خلاصه توضیحاتی در مورد آن‌ها ارائه می‌دهیم.

Training set^۴
Validation set^۵
Test set^۶

مجموعه داده بیماری دیابت

این مجموعه شامل اطلاعات مربوط به ۷۶۸ نفر (همگی زن با سن بالاتر از ۲۱ سال) است. در این مجموعه داده، ۸ مشخصه از افراد شامل موارد زیر گنجانده شده است [۱۴]:

1. Number of times pregnant
2. Plasma glucose concentration at 2 hours in an oral glucose tolerance test
3. Diastolic blood pressure (mm Hg)
4. Triceps skin fold thickness (mm)
5. 2-hour serum insulin (μ U/ml)
6. Body mass index (weight in kg/(height in m)²)
7. Diabetes pedigree function
8. Age (years)

مجموعه مذکور دارای دو برچسب است که نشان دهنده وجود یا عدم وجود دیابت در افراد تحت آزمایش است. تست دیابت در ۶۵/۱٪ از نمونه‌های موجود در این مجموعه منفی می‌باشد.

مطابق آنچه که در بخش ۳ توضیح داده شد، برای اعمال این مجموعه داده به شبکه عصبی پیشنهادی، ابتدا اطلاعات ۷۶۸ نفر به سه دسته تقسیم شد. بدین‌صورت که از ۳۸۴ نمونه به عنوان داده‌های آموزشی، از ۱۹۲ نمونه به عنوان داده‌های تأییدی و از ۱۹۲ نمونه باقی‌مانده به عنوان داده‌های آزمونی استفاده شد. برای اطمینان بیشتر از صحت عملکرد شبکه، مجموعه داده در سه نسخه مختلف و تحت اسامی دیابت ۱ تا ۳ تقسیم‌بندی شد. تنها تفاوت این سه نسخه در داده‌هایی است که برای مراحل آموزش، تأیید و آزمون بکار رفته‌اند.

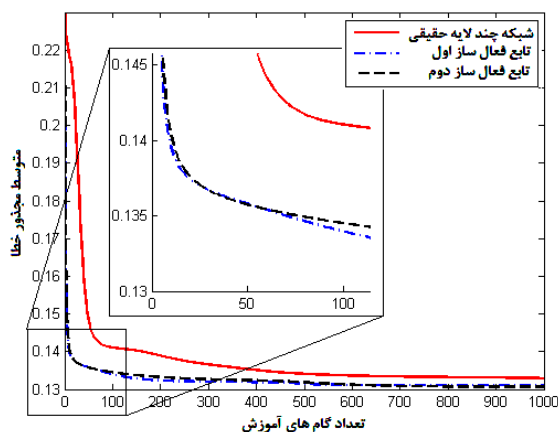
در مرحله آموزش شبکه، مقدار نرخ یادگیری سیستم (η) برابر با ۰/۱ انتخاب شد. همچنین قسمت‌های حقیقی و موهومی وزن‌های مختلط و بایاس‌های اولیه شبکه به صورت تصادفی و با توزیع یکنواخت در بازه $[-۰/۵, ۰/۵]$ قرار داده شدند. دو تابع فعال‌ساز پیشنهاد شده در مرجع [۱۲] نیز با استفاده از روابط اصلاح وزن استخراج شده در آن و مطابق روش پیشنهاد شده در بخش ۳ توسط داده‌های مذکور آموزش داده شدند و نتایج حاصل از آن‌ها برای مقایسه با توابع فعال‌ساز پیشنهادی مقاله مورد استفاده قرار گرفتند.

همچنین برای مقایسه عملکرد شبکه عصبی با مقادیر مختلط (با توابع فعال‌ساز پیشنهادی) با شبکه عصبی با مقادیر حقیقی، از یک شبکه پرسپترون چند لایه استاندارد (با مقادیر وزن حقیقی)

استفاده شد. برای منصفانه بودن مقایسه از نظر سرعت و دیگر مشخصات، سعی شد تا ساختار شبکه پرسپترون چند لایه به گونه‌ای انتخاب شود که تعداد پارامترهای آن (شامل وزن‌ها و بایاس‌ها) تقریباً با تعداد پارامترهای شبکه مختلط برابر باشد. لازم به ذکر است که هر یک از وزن‌ها و بایاس‌های مختلط، به صورت دو پارامتر در نظر گرفته شده‌اند زیرا دارای دو بخش حقیقی و موهومی هستند. با توجه به آنچه بیان شد، یک شبکه پرسپترون سه لایه استاندارد، با سه نورون در لایه مخفی انتخاب گردید. بدین ترتیب تعداد پارامترهای شبکه عصبی مختلط ۳۶ عدد و پارامترهای پرسپترون سه لایه ۳۵ عدد می‌باشد (شبکه‌ها دارای ۸ ورودی و ۲ خروجی و یک نورون بایاس می‌باشند).

در نمودار شکل ۵ مقدار متوسط مجذور خطای یادگیری سیستم در طی گام‌های آموزش، برای دیابت ۳ رسم شده است. ملاحظه می‌شود که سرعت همگرایی شبکه با مقادیر مختلط با هر دو تابع فعال‌ساز پیشنهادی، بیشتر از شبکه پرسپترون سه لایه استاندارد بوده و همچنین دارای متوسط مجذور خطای کمتری نیز می‌باشند. واضح است که سرعت همگرایی شبکه عصبی با مقادیر مختلط بسیار بیشتر از شبکه حقیقی سه لایه می‌باشد و شبکه عصبی چند لایه حقیقی به تعداد گام‌های بیشتری برای آموزش و رسیدن به حداقل مجذور خطا نیاز دارد.

برای مقایسه بهتر، نتایج تست شبکه بر روی مجموعه داده مربوط به دیابت برای شبکه‌های عصبی با مقادیر مختلط توسط توابع فعال‌ساز پیشنهادی و توابع فعال‌ساز موجود در مرجع [۱۲] و همچنین نتایج حاصل از شبکه پرسپترون سه لایه حقیقی در جدول ۲ ارائه شده‌اند. در این جدول، مقدار متوسط خطای دسته بندی و تعداد دفعات مورد نیاز برای رسیدن به حداقل خطای مرحله تأیید ذکر شده‌اند. نتایج موجود در جدول، حاصل متوسط‌گیری از نتایج ۲۰ بار شبیه‌سازی شبکه‌ها می‌باشند.



شکل ۵. مقدار متوسط مجذور خطای یادگیری سیستم در طی گام‌های آموزش برای دیابت ۳

دیده می‌شود که متوسط خطای دسته بندی برای هر دو توابع فعال‌ساز تقریباً با یکدیگر نزدیک بوده و خطای دسته‌بندی آن‌ها بهتر از توابع فعال‌ساز مرجع [۱۲] می‌باشد (به استثنای حالت مجموعه داده دیابت ۲ توسط تابع فعال‌ساز دوم). همچنین نتایج هر چهار تابع فعال‌ساز شبکه‌های عصبی با مقادیر مختلف، عملکرد بهتری از نظر متوسط خطای دسته‌بندی داشته‌اند.

از نظر تعداد گام‌های مورد نیاز جهت رسیدن به حداقل خطای مرحله تأیید، توابع فعال‌ساز پیشنهادی به وضوح عملکرد بسیار مناسبی از خود نشان داده‌اند. لذا با این تفاوت آشکار میان گام‌های مورد نیاز، به راحتی می‌توان دریافت که سرعت همگرایی این شبکه‌ها بسیار بیشتر از سه شبکه دیگر می‌باشد.

جدول ۲. نتایج شبیه‌سازی شبکه عصبی برای تشخیص دیابت

مجموعه داده	متوسط خطای دسته بندی			تعداد گام‌ها در حداقل خطای مرحله تأیید		
	فعال‌ساز اول	فعال‌ساز دوم	فعال‌ساز اول	فعال‌ساز دوم	فعال‌ساز اول	فعال‌ساز دوم
سه لایه حقیقی	فعال‌ساز اول [۱۲]	فعال‌ساز دوم [۱۲]	سه لایه حقیقی	فعال‌ساز اول [۱۲]	فعال‌ساز دوم [۱۲]	سه لایه حقیقی
دیابت ۱	۴۰.۴±۱.۶	۴۰.۷±۱.۴۱	±۱.۵۳	±۱.۰۸	±۱.۷۹	۴۴۷.۲۵
دیابت ۲	±۱.۰۷	۴۸.۲±۱	۴۵.۶±۱.۵۶	۴۶.۷±۱.۳	±۲.۰۸	۴۶۴.۷۵
دیابت ۳	±۱.۷۱	±۱.۷۷	۴۰.۶±۱.۴۶	±۱.۷۱	±۱.۸۴	۸۱.۷۵
	۴۰.۳	۴۰.۷۵	۴۴.۱۵	۴۱.۹	۵.۲۵	۶۳۷.۵

با توجه به نتایج بدست آمده از شبیه‌سازی، شبکه عصبی با تابع فعال‌ساز اول، در هر سه نسخه مجموعه داده، در حدود ۸۰٪ مواقع تشخیص صحیح داده است. در تابع فعال‌ساز ۲ نیز میزان موفقیت در بدترین حالت برابر با ۷۵٪ می‌باشد. ملاحظه می‌شود که توابع فعال‌ساز پیشنهادی عملکرد مناسب‌تری (هم از لحاظ متوسط خطای دسته بندی و هم از جهت تعداد گام‌های مورد نیاز جهت رسیدن به حداقل خطای مرحله تأیید) نسبت به شبکه سه لایه حقیقی دارند. برتری شبکه با توابع فعال‌ساز پیشنهادی، در تعداد گام‌های مورد نیاز برای رسیدن به حداقل خطای مرحله تأیید، به وضوح آشکار است.

مجموعه داده سرطان سینه

این مجموعه شامل اطلاعات مربوط به ۶۹۹ نفر است که ۶۵.۵٪ افراد این مجموعه داده، مبتلا به سرطان سینه نوع خوش‌خیم هستند. پارامترهای موجود در این مجموعه داده به شرح زیر است [۱۴]:

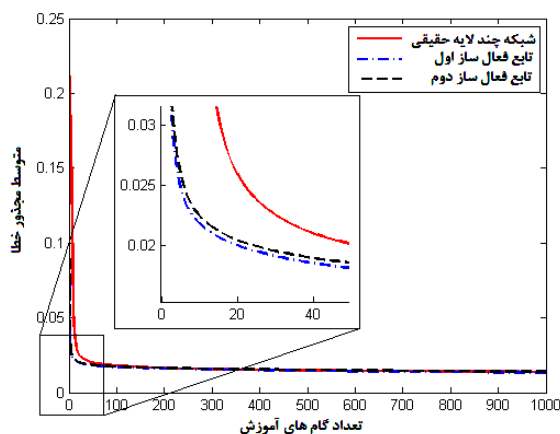
1. Clump thickness
2. Uniformity of cell size
3. Uniformity of cell shape
4. Marginal adhesion
5. Single epithelial cell size
6. Bare chromatin
7. Bland chromatin
8. Normal nucleoli
9. Mitoses

مجموعه مذکور دارای دو برجسب است که نشان دهنده خوش‌خیم یا بدخیم بودن سرطان سینه در فرد است.

در این قسمت نیز مطابق با مطالب ذکر شده در بخش ۳ این مقاله عمل کرده و اطلاعات ۶۹۹ نمونه به سه گروه دسته بندی شد. ۳۵۰ نمونه به عنوان داده‌های آموزشی، ۱۷۵ نمونه به عنوان داده‌ها تأییدی و ۱۷۴ نمونه باقی‌مانده به عنوان داده‌های آزمونی استفاده شدند. در اینجا نیز

همانند بخش قبلی، برای اطمینان از صحت عملکرد شبکه عصبی، مجموعه داده‌ها در سه نسخه مختلف و تحت اسامی سرطان ۱-۳ تقسیم‌بندی شدند و به سه شبکه عصبی اعمال گردیدند. در نمودار شکل ۶ مقدار متوسط مجذور خطای یادگیری در طی گام‌های آموزش برای مجموعه داده سرطان ۳ رسم شده است. در این نمودار نیز سرعت همگرایی شبکه مختلط با دو تابع فعال‌ساز پیشنهادی، بیشتر از شبکه پرسپترون سه لایه با مقادیر حقیقی است.

در جدول ۳ نتایج تست شبکه بر روی مجموعه داده مربوط به سرطان سینه برای هر دو تابع فعال‌ساز پیشنهاد شده و توابع فعال‌ساز مرجع [۱۲] و همچنین شبکه پرسپترون سه لایه با وزن‌های حقیقی، ارائه شده است. در این جدول، مقدار متوسط خطای دسته بندی و تعداد دفعات مورد نیاز برای رسیدن به حداقل خطای مرحله تأیید ذکر شده‌اند. همانند قسمت قبلی، نتایج، حاصل متوسط‌گیری از نتایج ۲۰ بار شبیه‌سازی شبکه است.



شکل ۶. مقدار متوسط مجذور خطای یادگیری سیستم در طی گام‌های آموزش برای سرطان ۳

برای این مجموعه داده نیز مشاهده می‌شود که تابع فعال‌ساز اول تقریباً در اکثر موارد متوسط خطای دسته بندی کمتری دارد. اما عملکرد فعال‌ساز دوم در برخی موارد بدتر از توابع دیگر بوده اما این اختلاف، در حد زیادی نبوده و تقریباً می‌توان عملکرد آن‌ها را شبیه به هم دانست.

اما برای این مجموعه داده همانند مجموعه داده دیابت، شبکه عصبی با مقادیر مختلط با توابع فعال‌ساز اول و دوم، به تعداد گام‌های کمتری جهت رسیدن به حداقل خطای مرحله تأیید نیاز دارند و سرعت همگرایی بیشتری نسبت به سه شبکه دیگر دارند. با توجه به نتایج بدست آمده از

شبیه‌سازی، شبکه عصبی با تابع فعال‌ساز اول و دوم، در بدترین حالت، در حدود ۹۵٪ مواقع تشخیص صحیح داده است.

جدول ۳. نتایج شبیه‌سازی شبکه عصبی برای تشخیص سرطان سینه

متوسط خطای دسته بندی					تعداد گام‌ها در حداقل خطای مرحله تأیید				
مجموعه داده	فعال‌ساز اول	فعال‌ساز دوم	فعال‌ساز اول [۱۲]	فعال‌ساز دوم [۱۲]	سه لایه حقیقی	فعال‌ساز اول	فعال‌ساز دوم	فعال‌ساز اول [۱۲]	سه لایه حقیقی
سرطان ۱	7 ± 0.32	8.1 ± 0.85	7 ± 0.47	7 ± 0.58	7.95 ± 0.22	۱۶.۲۵	۵.۲۵	۲۸.۲۵	۲۱
سرطان ۲	7.65 ± 0.67	7.9 ± 0.44	7.35 ± 0.81	8.12 ± 0.71	6.75 ± 0.44	۱۵.۲۵	۲۲.۲۵	۲۲.۲۵	۱۶
سرطان ۳	7.1 ± 0.3	8.35 ± 1.08	7.48 ± 0.27	7.2 ± 0.46	8 ± 0.13	۱۵.۷۵	۵	۲۸.۵	۲۰.۲۵

نتیجه‌گیری

در این مقاله به کاربرد شبکه‌های عصبی مصنوعی با مقادیر مختلط در مسائل دسته‌بندی با مقادیر حقیقی پرداخته شد. نشان داده شد که این شبکه‌ها نقص اصلی پرسپترون دو لایه‌ی عادی، در حل مسائل جداناپذیر خطی را ندارند. لذا یکی از مهم‌ترین مزایای این نوع از شبکه‌های عصبی، دو لایه بودن شبکه و عدم نیاز به لایه‌های پنهان در ساختار شبکه عصبی است که این ویژگی خود موجب سادگی ساختار شبکه و در نتیجه کاهش میزان محاسبات لازم می‌شود.

دو تابع فعال‌ساز جدید برای شبکه عصبی مختلط پیشنهاد شد. روابط اصلاح وزن برای هرکدام از توابع پیشنهادی استخراج و روند آموزش و تست شبکه عصبی شرح داده شد. برای بررسی عملکرد ساختارهای پیشنهادی در حوزه تشخیص پزشکی، مجموعه داده‌های مربوط به بیماری دیابت و سرطان سینه به شبکه اعمال گردیدند. جهت مقایسه عملکرد شبکه با توابع فعال‌ساز پیشنهادی، داده‌ها به یک شبکه چند لایه پرسپترون استاندارد با مقادیر حقیقی نیز اعمال گردیدند. نتایج نشان داد که شبکه مختلط با هر دو تابع فعال‌ساز، بسیار سریع‌تر از شبکه چند لایه پرسپترون استاندارد به مقدار مینیمم متوسط مجذور خطا همگرا شده و مینیمم متوسط مجذور خطای کمتری نیز دارند. شبکه چند لایه پرسپترون به گونه‌ای انتخاب گردید تا تقریباً پارامترهای شبکه‌ها و در نتیجه محاسبات مورد نیاز تقریباً برابر باشند. میزان حداقل گام‌های مورد نیاز، جهت رسیدن به حداقل خطای مرحله‌ی تأیید، در شبکه پرسپترون سه لایه، تفاوت قابل ملاحظه‌ای با تعداد گام‌ها در شبکه‌های پیشنهادی دارد که این امر سرعت بالای همگرایی این شبکه‌ها را نشان می‌دهد.

برای مقایسه توابع فعال‌ساز پیشنهادی، از دو تابع فعال‌ساز پیشنهادی دیگر در یکی از مراجع استفاده شد. نتایج نشان داد که تابع فعال‌ساز اول در همه موارد عملکرد بهتری نسبت به توابع فعال‌ساز دیگر دارد. تابع فعال‌ساز دوم نیز در قسمت‌هایی عملکرد بهتری و در جاهایی عملکرد ضعیف‌تری از نظر خطای دسته بندی از خود نشان داد. می‌توان یکی از دلایل عملکرد بهتر توابع فعال‌ساز پیشنهادی نسبت به توابع مرجع مذکور را، استفاده از چهار ناحیه اشباع توسط هر دو تابع فعال‌ساز پیشنهادی دانست، زیرا یکی از توابع فعال‌ساز مرجع مذکور، تنها از سه ناحیه اشباع بهره می‌برد.

البته با توجه به کم بودن میزان اختلاف میان متوسط خطای دسته‌بندی، می‌توان عملکرد هر چهار را، از این نظر تقریباً یکسان دانست. اما از نظر تعداد گام‌های مورد نیاز جهت رسیدن به حداقل خطای مرحله تأیید، توابع فعال‌ساز پیشنهادی عملکرد بسیار مطلوبی را نشان دادند. تفاوت این گام‌ها در مجموعه داده دیابت بسیار واضح‌تر بوده و عملکرد سریع‌تر توابع پیشنهادی نسبت به توابع مرجع مذکور را نشان می‌دهد.

نتایج شبیه‌سازی نشان داد که شبکه عصبی مختلط با ساختارهای پیشنهادی، در تشخیص بیماری دیابت در ۸۰٪ موارد به درستی عمل تشخیص را انجام می‌دهد. همچنین در مورد تشخیص سرطان سینه، درصد موفقیتی برابر با ۹۵٪ بدست آمد.

مراجع

- [1] G. M. Georgiou and C. Koutsougeras, "Complex domain backpropagation," *Circuits and Systems II: Analog and Digital Signal Processing*, IEEE Transactions on, 1992 vol. 39, pp. 330–334.
- [2] N. Benvenuto and F. Piazza, "On the complex backpropagation algorithm," *Signal Processing*, IEEE Transactions on, 1992, vol. 40, pp. 967–969.
- [3] T. Nitta, "An extension of the back-propagation algorithm to complex numbers," *Neural Networks*, 1997, vol. 10, pp. 1391–1415.
- [4] A. Hirose, *Complex-valued neural networks: theories and applications*, 2004 vol. 5: World Scientific Publishing Company Incorporated.
- [5] T. Nitta, *Complex-valued neural networks: Utilizing high-dimensional parameters*: Information Science Reference, 2009.
- [6] A. S. Gangal, P. K. Kalra, and D. S. Chauhan, "Performance evaluation of complex valued neural networks using various error functions," *Enformatika*, 2007 vol. 23, pp. 27–32.
- [7] B. Karlik and A. V. Olgac, "Performance analysis of various activation functions in generalized mlp architectures of neural networks," *International*

- Journal of Artificial Intelligence and Expert Systems, 2010, vol. 1, pp. 111–122.
- [8] E. Kreyszig, Advanced Engineering Mathematics: John Wiley & Sons, 2006.
- [9] T. Kim and T. Adali, "Fully complex multi-layer perceptron network for nonlinear signal processing," The Journal of VLSI Signal Processing, 2002, vol. 32, pp. 29–43.
- [10] H. Leung and S. Haykin, "The complex backpropagation algorithm," Signal Processing, IEEE Transactions on, 1991, vol. 39, pp. 2101–2104.
- [11] H. E. Michel and A. A. S. Awwal, "Enhanced artificial neural networks using complex numbers," presented at the Neural Networks, 1999. IJCNN'99. International Joint Conference on, 1999.
- [12] F. Amin and K. Murase, "Single-layered complex-valued neural network for real-valued classification problems," Neurocomputing, 2009, vol. 72, pp. 945–955.
- [13] l.prechelt. "a set of benchmarks and benchmarking rules for neural network training algorithms," university of karlsruhe, germany. [online]. available: <ftp.ira.uka.de/pub/neuron/proben1.tar.gz>
- [14] "Machine learning repository-center for machine learning and intelligent systems." [Online]. Available: <http://archive.ics.uci.edu/ml/datasets.html>

Archive of SID