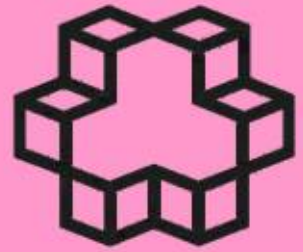




دانشگاه صنعتی خواجه نصیرالدین طوسی
هفتمین کنفرانس سیستم‌های هوشمند



7th Conference of Intelligence Systems
University of K.N.T.U

۲۹ و ۳۰ آذر و اول دی ماه ۱۳۸۴

نویسندگان محترم

با اهداء سلام

احتراما پيرو قبولی مقاله شما با عنوان:

یک الگوریتم آموزش جدید برای شبکه عصبی MLP در کاربردهای
دسته‌بندی

تقاضا می‌شود آخرین نسخه اصلاح شده را جهت چاپ و درج در کتابچه
کنفرانس با تکمیل فرم ثبت‌نام تا تاریخ ۱۳۸۴/۰۸/۳۰ از طریق پست
الکترونیکی info@cis2005.info ارسال نمایید.

- فقط مقالاتی به چاپ خواهند رسید که بصورت شفاهی در زمان برگزاری
کنفرانس ارائه گردند.
- برای دریافت فرم ثبت نام و اطلاع از نحوه برگزاری کنفرانس به سایت
اینترنتی www.cis2005.info مراجعه نمایید.

محمد تشنه‌لب

دبیر هفتمین کنفرانس

سیستم‌های هوشمند

تهران، تقاطع سید خندان و شریعتی - نبش شرکت پست، دانشکده برق

دانشگاه صنعتی خواجه نصیرالدین طوسی

صندوق پستی ۱۳۵۹-۱۶۳۱۵ تلفکس: ۸۸۴۶۲۳۲۹

یک الگوریتم آموزش جدید برای شبکه عصبی MLP در کاربردهای دسته‌بندی

کامران کاظمی	مهدی سعادت‌مند طرزجان	محمد تشنه‌لب	حمید ابریشمی مقدم
دانشگاه صنعتی خواجه نصیرالدین	دانشگاه تربیت مدرس،	دانشگاه صنعتی خواجه نصیرالدین	دانشگاه صنعتی خواجه نصیرالدین
طوسی، تهران، ایران	تهران، ایران	طوسی، تهران، ایران	طوسی، تهران، ایران
ka_kazemi@eetd.kntu.ac.ir	saadatmand@kiaeee.org	teshnehlab@eetd.knta.ac.ir	moghadam@eetd.kntu.ac.ir

چکیده: مرحله آموزش وزن‌ها، مهمترین بخش در پیاده‌سازی یک شبکه عصبی می‌باشد. در این مقاله یک الگوریتم آموزشی جدید برای شبکه عصبی MLP در کاربردهای دسته‌بندی ارائه شده است. الگوریتم پیشنهادی با ایجاد مصالحه‌ای بین معیار دسته‌بندی و معیار آموزش شبکه عملاً خطای شبکه را به شکل واقعی‌تری محاسبه می‌کند. از الگوریتم پیشنهادی برای دسته‌بندی پیکسل‌های یک تصویر ماهواره‌ای چند-باندی استفاده شده است. عملکرد الگوریتم بر روی این مسأله با الگوریتم‌های آموزشی BP و ELS و الگوریتم کلاسیک ML مقایسه شده است. شبیه‌سازی‌ها بیانگر آن است که دقت الگوریتم پیشنهادی از تمام روش‌های فوق بهتر می‌باشد.

کلمات کلیدی: دسته‌بندی داده‌ها، شبکه عصبی MLP، الگوریتم آموزشی BP، تصاویر سنجنش از دور

۱- مقدمه

دسته‌بندی (classification)، تخمین توابع (function approximation) و شناسایی سیستم (system identification) از جمله کاربردهای شبکه‌های عصبی مصنوعی می‌باشند. به دلیل خصوصیتی چون آموزش‌پذیری و غیرخطی بودن، دسته‌بندی یکی از مهمترین کاربردهای شبکه‌های عصبی می‌باشد [۱]. علاوه بر شبکه‌های عصبی در مقایسه با الگوریتم‌های پارامتری از قبیل حداکثر شباهت، ML (maximum likelihood)، و بیشترین احتمال پیشین، MAP (maximum a posteriori)، در زمره الگوریتم‌های دسته‌بندی غیرپارامتری بشمار می‌آیند. آموزش وزن‌ها مهمترین بخش در پیاده‌سازی یک شبکه عصبی است. الگوریتم‌های آموزشی خود به دو دسته تقسیم می‌شوند: (۱) با مربی و (۲) بدون مربی. برخلاف الگوریتم‌های آموزش بدون مربی که در حقیقت داده‌های ورودی را بر اساس شباهت‌هایشان خوشه‌بندی می‌کنند [۲]، الگوریتم‌های آموزش با مربی مجموعه‌ای از داده‌های آموزشی را بکار می‌گیرند. پاسخ مطلوب شبکه برای هر یک از آنها از قبل مشخص می‌باشد. در بین شبکه‌های متعددی که در کاربردهای دسته‌بندی مورد استفاده قرار گرفته‌اند از قبیل RBF (radial basis functions) [۳]، SVM (support vector machine) [۴] و HLVQ (hybrid learning vector quantization) [۵]؛ شبکه عصبی MLP (multi-layer perceptron) [۶] [۷] متداولترین و پرکاربردترین شبکه است. معمولاً هنگامی که از MLP به عنوان یک دسته‌بندی کننده استفاده می‌شود، متناظر با هر کلاس تعلق، یک نرون در لایه خروجی آن در نظر گرفته می‌شود. MLP داده ورودی را به کلاس تعلق نسبت می‌دهد که نرون متناظر با آن بیشترین خروجی را داشته باشد (معیار دسته‌بندی). تاکنون الگوریتم‌های آموزشی متعددی از قبیل پس‌انتشار خطا، BP (back propagation) [۸]، الگوریتم‌های برپایه کمترین مربعات (LSB) [۹]، کمترین مربعات توسعه‌یافته (ELS) [۱۰] و روش لونیبرگ مارکوویت (LM) [۱۱] برای آموزش MLP پیشنهاد گردیده است. در تمام الگوریتم‌های فوق، معیار آموزش وزن‌ها با معیار دسته‌بندی متفاوت است. در این مقاله نشان داده شده است، که چگونه دوگانگی موجود بین معیار آموزش شبکه و معیار دسته‌بندی می‌تواند سبب کاهش قابلیت انعطاف شبکه گردد. با هدف رفع دوگانگی فوق، ما یک الگوریتم آموزشی جدید برای

شبکه MLP ارائه نموده‌ایم. الگوریتم پیشنهادی با تنظیم وفقی خروجی مطلوب شبکه، سعی دارد معیار آموزش را تا حد امکان به معیار دسته‌بندی نزدیک نماید. از الگوریتم آموزشی پیشنهادی برای دسته‌بندی داده‌های یک تصویر ماهواره‌ای چند بانندی، استفاده شده است. همچنین عملکرد الگوریتم پیشنهادی با الگوریتم‌های آموزشی BP و ELS و روش کلاسیک ML مقایسه شده است. شبیه‌سازی‌ها بیانگر آن است که دقت الگوریتم پیشنهادی در دسته‌بندی داده‌ها، از تمام روش‌های فوق بهتر بوده است. در ادامه مقاله، در بخش بعد، روش متداول در آموزش شبکه‌های عصبی بررسی می‌شود. بخش سوم به شرح الگوریتم پیشنهادی می‌پردازد. شبیه‌سازی‌ها در بخش چهارم آورده شده است و در نهایت، بخش پنجم به جمع‌بندی مقاله اختصاص یافته است.

۲- الگوریتم‌های آموزشی متداول

فرض کنید تعداد N نمونه آموزشی در M کلاس تعلق مختلف در دست باشد. در روش متداول آموزش شبکه‌های عصبی به منظور دسته‌بندی داده‌ها، برای هر داده آموزشی $\bar{x}_i (i=1,2,\dots,N)$ ، خروجی مطلوب شبکه $\bar{y}_i (i=1,2,\dots,N)$ به صورت ذیل تعریف می‌شود:

$$\bar{x}_i \in \omega_k \Leftrightarrow y_{i,j} = \begin{cases} \eta_{mx}, & j = k \\ \eta_{mn}, & \text{Otherwise} \end{cases} \quad (1)$$

که در آن ω_k ، $y_{i,j}$ ، η_{mx} و η_{mn} به ترتیب عبارتند از کلاس تعلق k عنصر j خروجی مطلوب \bar{y}_i ، مقدار خروجی مطلوب برای نرون متناظر با کلاس داده ورودی و مقدار خروجی مطلوب برای سایر نرون‌ها. به عنوان مثال چنانچه از توابع فعالیت MLP از نوع (\cdot) انتخاب شود، مقادیر مناسب برای η_{mn} و η_{mx} به ترتیب 0.8 و -0.8 است [۸]. می‌توان گفت اکثر قریب به اتفاق الگوریتم‌های آموزشی، بر مبنای کمینه کردن یک معیار بر روی نمونه‌های آموزشی طراحی شده‌اند. به عنوان مثال، BP مطابق معادله (۲) بر اساس معیار کمترین مربعات خطا، LSE (least square errors)، طراحی شده است.

$$E = \frac{1}{2} \sum_{i=1}^N \|\bar{e}_i\|^2, \quad e_i = \bar{\psi}_i - \bar{y}_i \quad (2)$$

که در آن، ψ_i و e_i به ترتیب خروجی شبکه به ازای داده آموزشی i و خطای آن است. در هر مرحله از BP، وزن‌های شبکه، با پس‌انتشار خطای خروجی در لایه‌های پنهان، بهنگام می‌گردند. پس از اتمام مرحله آموزش، کلاس تعلق داده ورودی \bar{x} مطابق معادله ذیل تعیین می‌شود.

$$\bar{x} \in \omega_p, \quad p = \text{index} \left(\max_{k=1}^M (\zeta_k) \right) \quad (3)$$

که در آن ζ_k خروجی نرون k (در لایه خروجی) به ازای داده ورودی \bar{x} است. همانطور که مشاهده می‌شود، در همه الگوریتم‌های متداول برای آموزش MLP مانند BP، LS، ELS و LM، معیار آموزش شبکه (معادله ۲) با معیار دسته‌بندی داده‌ها (معادله ۳) متفاوت است. دوگانگی فوق می‌تواند سبب کاهش قابلیت انعطاف شبکه در فرآیند آموزش شود. به عنوان مثال، چنانچه داده ورودی متعلق به کلاس اول از بین پنج کلاس و خروجی شبکه نیز $\bar{\psi} = [\zeta_k] = [0.8, 0.7, 0.0, -0.8]^T$ باشد، آنگاه طبق معیار دسته‌بندی (معادله ۳)، شبکه داده ورودی را بدرستی دسته‌بندی می‌کند اما طبق LSE (معادله ۲)، داده مذکور دارای خطای ذیل است.

$$\|\bar{e}\| = \|[0.0, 1.5, 0.8, 0.0]^T\| = 1.7. \quad (4)$$

اکنون، فرض کنید که بردار خروجی شبکه برای همان داده ورودی $\bar{\psi} = [-0.7, 0.0, -0.8, -0.8]^T$ باشد. در این حالت، داده به غلط، به کلاس دوم نسبت داده خواهد شد، در حالی که مطابق با معیار LSE، خطای آن باز هم برابر با $\|\bar{e}\| = \|[-1.5, 0.8, 0.0, 0.0]^T\| = 1.7$ است. به عبارت دیگر، با اینکه شبکه در حالت اول برخلاف حالت دوم داده را بدرستی دسته‌بندی نموده است، اما در هر دو حالت فوق، معیار LSE خطای یکسانی را به داده نسبت می‌دهد. همانطور که مشاهده می‌شود، متفاوت بودن معیار آموزش (معادله ۲) و معیار دسته‌بندی (معادله ۳) سبب می‌شود که گاهی به داده‌ای که بدرستی دسته‌بندی شده است، خطای بزرگی نسبت داده شود.

در این مقاله، با هدف حل مشکل فوق، الگوریتمی جدید برپایه تنظیم وفقی خروجی مطلوب شبکه در طول فرآیند آموزش ارائه شده است. در حقیقت، زمانی که هدف ما استفاده از شبکه عصبی برای دسته‌بندی داده‌ها است، در نظر گرفتن یک خروجی

مطلوب ثابت برای هر داده آموزشی الزامی نیست. زیرا تنها شرط دسته‌بندی صحیح داده ورودی آن است که در بین نرون‌های لایه آخر شبکه، نرون مربوط به کلاس صحیح، دارای بیشترین خروجی باشد. به بیان دیگر، در کاربردهای دسته‌بندی، برخلاف شناسایی سیستم و تخمین توابع، برای هر داده بینهایت خروجی معتبر وجود دارد که همه آنها در معادله (۳) صدق می‌کنند. لذا در این نوع کاربردها، لزومی ندارد که طی فرآیند آموزش، برای هر داده تنها یک خروجی مطلوب ثابت در نظر گرفته شود. الگوریتم آموزشی پیشنهادی، خروجی مطلوب شبکه برای هر داده، براساس خروجی شبکه و کلاس تعلق داده ورودی به صورت وقتی تنظیم می‌شود.

۳- الگوریتم آموزش پیشنهادی

همانطور که در بخش قبل شرح داده شد، دوگانگی بین معیار آموزش و معیار دسته‌بندی سبب می‌شود که خطای داده‌های ورودی به درستی محاسبه نشود. همچنین نشان داده شد که در کاربردهای دسته‌بندی، می‌توان برای یک داده آموزشی، بینهایت خروجی مطلوب در نظر گرفت. در الگوریتم پیشنهادی از این خصوصیت برای تصحیح معیار آموزش استفاده شده است. در حقیقت، الگوریتم پیشنهادی خروجی‌های مطلوب را در طول فرآیند آموزش چنان تغییر می‌دهد که خطای داده‌ها به شکل واقعی‌تری محاسبه گردد.

هدف ما این است که به داده‌هایی که شبکه آنها را بر اساس معیار دسته‌بندی بدرستی دسته‌بندی می‌کند، در مقایسه با سایر داده‌های آموزشی، خطای بسیار کوچکتری نسبت داده شود. طبق معیار LSE، خروجی شبکه برای داده \vec{x} ، تنها زمانی بدون خطا است که در معادله ذیل صدق کند،

$$\zeta_C = \zeta_k + \eta_{mx} - \eta_{mn}, \quad k = 1, 2, \dots, M, \quad k \neq C, \quad (5)$$

که در آن C بیانگر کلاس تعلق \vec{x} است. در حالی که در معیار دسته‌بندی برقراری نامعادله ذیل شرط لازم و کافی بدون خطا بودن خروجی شبکه است.

$$\zeta_C > \zeta_k, \quad k = 1, 2, \dots, M, \quad k \neq C \quad (6)$$

در روش پیشنهادی، خروجی شبکه باید در نامعادله (۷) صدق نماید تا خطای آن صفر گردد،

$$\zeta_C > \zeta_k + \alpha, \quad k = 1, 2, \dots, M, \quad k \neq C, \quad (7)$$

که در آن α عددی مثبت و در بازه $[0, 1]$ است. با توجه به معادله (۱) به سادگی می‌توان نشان داد که به ازای مقادیر یک و صفر برای α ، نامعادله (۷) به ترتیب به معادله‌های (۵) و (۶) ساده می‌شود. در حقیقت معادله (۷) مصالحه‌ای بین معیار دسته‌بندی و معیار LSE ایجاد می‌کند. پارامتر α تعیین کننده میزان آزادی شبکه در فرآیند آموزش است. هرچه α به یک نزدیکتر باشد، معیار آموزش سختگیرانه‌تر و هرچه به صفر نزدیکتر باشد ساده‌تر است. بر این اساس، خروجی مطلوب شبکه برای داده آموزشی \vec{x} در هر مرحله از فرآیند آموزش، به صورت ذیل محاسبه می‌شود،

$$y_k = \begin{cases} \eta_{mx}, & k = C \\ \eta_{mn}, & \eta_{mx} - \alpha < \zeta_k, \quad k = 1, 2, \dots, M \\ \zeta_k, & \text{Otherwise} \end{cases} \quad (8)$$

به این ترتیب با توجه به معادله (۲)، خطای خروجی شبکه برابر است با،

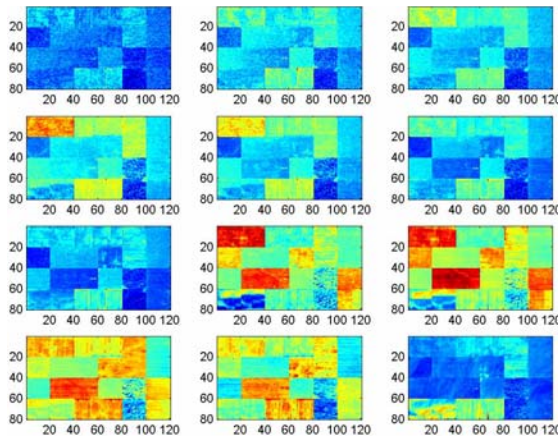
$$e_k = \begin{cases} \zeta_k - \eta_{mx}, & k = C \\ \zeta_k - \eta_{mn}, & \eta_{mx} - \alpha < \zeta_k, \quad k = 1, 2, \dots, M \\ 0, & \text{Otherwise} \end{cases} \quad (9)$$

همانطور که مشاهده می‌شود، در معیار پیشنهادی، خطای نرون متناظر با کلاس تعلق داده ورودی و نرون‌هایی که خروجی آنها بیش از اندازه بزرگ است (بیشتر از $\eta_{mx} - \alpha$ می‌باشد) مشابه با معیار LSE محاسبه می‌شود، در حالی که خطای سایر نرون‌ها مشابه با معیار دسته‌بندی صفر است. بنابراین الگوریتم آموزش پیشنهادی را می‌توان چنین خلاصه نمود که در هر مرحله از فرآیند آموزش در ابتدا خطای داده ورودی مطابق با معادله (۹) محاسبه می‌گردد. سپس وزن‌های شبکه با استفاده از خطای محاسبه شده در لایه خروجی و بر اساس روش BP بهنگام می‌گردند. مزیت الگوریتم پیشنهادی که آن را Extended BP (EBP) می‌نامیم در این است که به نرون‌هایی که (متناظر کلاس تعلق داده ورودی نبوده) خروجی آنها به اندازه کافی کوچک است، خطای صفر نسبت داده می‌شود. به این ترتیب، می‌توان انتشار داشت مه خطای شبکه به شکل واقعی‌تری محاسبه شود. به عنوان

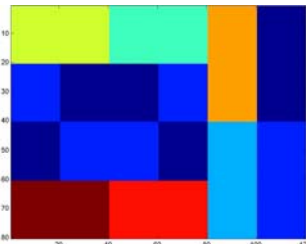
نمونه، خطای شبکه در مثالی که در بخش ۲ مطرح گردید، بر اساس معیار پیشنهادی، با $\alpha = 0.1$ در حالت اول $\|\bar{e}\| = 0.0$ و در حالت دوم برابر با $\|\bar{e}\| = \left\| \begin{bmatrix} -1.5, 0.0, 0.0, 0.0 \end{bmatrix} \right\| = 1.5$ است.

۴- شبیه‌سازی

برای بررسی عملکرد الگوریتم پیشنهادی، از یک تصویر ماهواره‌ای به ابعاد 120×80 که در دوازده باند گرفته شده است، استفاده نموده‌ایم. تعداد سطوح خاکستری در هر باند از تصویر مذکور، از ۲۵۶ سطح تشکیل شده است و تصویر شامل ۸ کلاس می‌باشد. کلاس‌های ۱ و ۲ هر کدام شامل ۲۴۰۰ پیکسل و سایر کلاس‌ها دارای ۸۰۰ پیکسل می‌باشند. در شکل‌های (۱) و (۲) به ترتیب سطوح خاکستری تصویر ماهواره‌ای (در ۱۲ باند) و ۸ کلاس آن، نشان داده شده است. از دسته‌بندی کلاس‌های تصویر ماهواره‌ای، به عنوان معیاری برای ارزیابی الگوریتم آموزشی پیشنهادی (EBP) استفاده شده است. در این مقاله عملکرد EBP با الگوریتم‌های آموزشی پس‌انتشار خطا (BP) [۸] و کمترین مربعات توسعه‌یافته (ELS) [۹] مقایسه شده است. همچنین



شکل ۱: سطوح خاکستری ۱۲ باند تصویر ماهواره‌ای

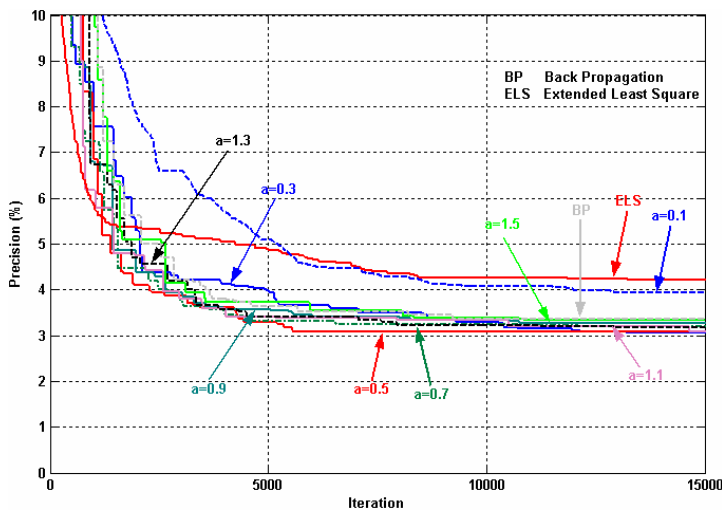


شکل ۲: هشت کلاس موجود در تصاویر ماهواره‌ای

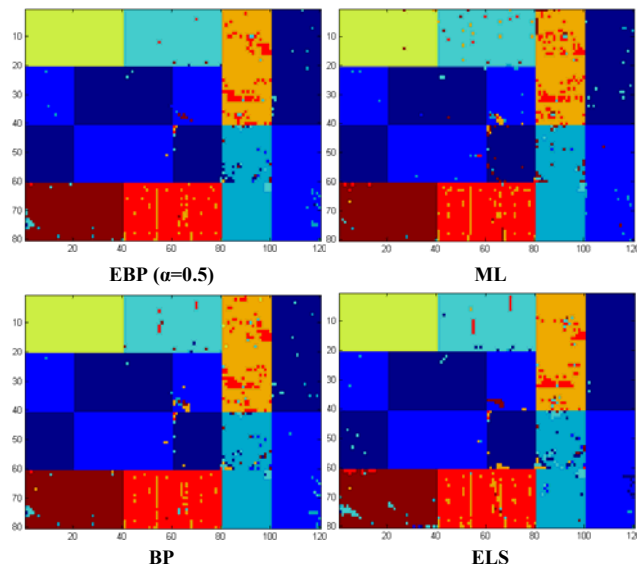
نتایج حاصل از الگوریتم با روش کلاسیک بیشترین شباهت (ML) نیز مقایسه گردیده است. برای آموزش شبکه عصبی صرف نظر از روش آموزش، از ۸۰۰ نمونه آموزشی (از هر کلاس ۱۰۰ نمونه) یکسان استفاده شده است. نمونه‌های آموزشی به صورت تصادفی انتخاب شده‌اند. برای ارزیابی عملکرد شبکه در طول فرآیند آموزش، از معیار خطای متوسط دسته‌بندی استفاده گردیده است. به این ترتیب که پس از هر مرحله آموزش، کلیه پیکسل‌های تصویر ماهواره‌ای، توسط شبکه دسته‌بندی گردیده و خطای متوسط دسته‌بندی در کلیه کلاس‌ها، در مقایسه با پاسخ مطلوب (شکل ۲) محاسبه می‌شود. بعلاوه، وزن‌های اولیه شبکه در کلیه شبیه‌سازی‌ها یکسان انتخاب شده‌اند. در شکل (۳) منحنی‌های تغییرات خطای متوسط دسته‌بندی در طول فرآیند آموزش شبکه با روش پیشنهادی (به ازای α های مختلف) در مقایسه با روش‌های BP و ELS نشان داده شده است. همانطور که در این شکل مشاهده می‌شود، مقدار α بر سرعت آموزش شبکه و دقت نهایی آن مؤثر است. به طوری که سرعت آموزش به ازای $\alpha = 0.5$ بیشینه گردیده است در حالی که بیشترین دقت مربوط به $\alpha = 0.3$ می‌باشد. با توجه به اینکه دقت نهایی شبکه با $\alpha = 0.3$ در مقایسه با $\alpha = 0.5$ تنها به اندازه ۰/۰۳٪ بهتر است، لذا $\alpha = 0.5$ به عنوان بهترین پاسخ انتخاب می‌شود.

با مراجعه به شکل (۳) می‌توان بروشنی مشاهده نمود که دقت نهایی و سرعت EBP به ازای تمام مقادیر α به استثنای $\alpha = 0.1$ ، بهتر از BP است. همچنین با اینکه در آغاز فرآیند آموزش، ELS خطا را با سرعت بیشتری نسبت به روش پیشنهادی (با $\alpha = 0.5$) کاهش داده است، اما این برتری تنها تا مرحله ۱۰۰۰ فرآیند آموزش حفظ گردیده است و پس از آن الگوریتم پیشنهادی از ELS پیشی گرفته است. بعلاوه، ELS برخلاف روش پیشنهادی در یک مینیمم محلی نه چندان مناسب بدام افتاده است. لذا دقت نهایی EBP از ELS بهتر می‌باشد.

در جدول (۱) دقت نهایی شبکه پس از ۱۵۰۰۰ مرحله به تفکیک کلاس‌ها نشان داده شده است. همانطور که مشاهده می‌شود، دقت نهایی EBP با $\alpha = 0.5$ در مقایسه با BP، ELS و ML به ترتیب بیش از ۰/۲۵٪، ۱/۱۲٪ و ۰/۸۱٪ بهبود یافته است که بیانگر کارایی الگوریتم پیشنهادی است. بعلاوه، الگوریتم پیشنهادی به ازای $\alpha = 0.3, 0.5, 1.1$ تقریباً پاسخ‌های یکسانی ارائه داده است که بیانگر توانایی الگوریتم در فرار از مینیمم‌های محلی و حساسیت کم الگوریتم نسبت به تغییرات α است. در شکل (۴) نتایج دسته‌بندی حاصل از الگوریتم پیشنهادی با $\alpha = 0.5$ مختلف در مقایسه با الگوریتم‌های BP، ELS و ML نشان داده شده



شکل ۳: منحنی‌های تغییرات خطای متوسط دسته‌بندی در طول آموزش شبکه با روش پیشنهادی به ازای α های مختلف در مقایسه با روش‌های BP و ELS



شکل ۴: نتایج دسته‌بندی حاصل از الگوریتم پیشنهادی در مقایسه با الگوریتم‌های BP، ML، ELS و EBP.

الگوریتم‌های BP و ELS مقایسه شده است. شبیه‌سازی‌ها بیانگر دقت و سرعت بیشتر الگوریتم پیشنهادی در مقایسه با سایر الگوریتم‌ها است. EBP با $\alpha = 0.5$ بهترین پاسخ‌ها را ارائه می‌دهد. دقت EBP در مقایسه با BP، ELS و ML به ترتیب، بیش از ۲۵٪، ۱۲٪ و ۸۱٪ بهبود داشته است که این خود بیانگر کارایی الگوریتم پیشنهادی است. همچنین سرعت EBP در مقایسه با سایر الگوریتم‌ها بهبود یافته است.

مراجع

- [1] S. Teodoridis, *Pattern recognition*. Academic Press, San Diego, 1999.
 [2] T. Kohonen, *Self-Organization and associative Memory*. Berlin: Springer-Verlag, 3rd ed, 1989.

است. همانطور که مشاهده می‌شود، پاسخ الگوریتم پیشنهادی به مراتب بهتر از سه الگوریتم دیگر است. در شکل (۵) منحنی دقت نهایی الگوریتم پیشنهادی بر حسب α نشان داده شده است. نکته جالب در منحنی مذکور تقارن موجود حول $\alpha = 0.8$ است. به طوری که دقت نهایی الگوریتم به ازای $\alpha = 0.5$ و $\alpha = 1.1$ تقریباً یکسان است. شاید بتوان این پدیده را با توجه به اینکه حد بالای خطا در معادله (۸)، 0.8 است چنین توضیح داد که مطابق معادله (۸)، $\alpha = 0.8$ معادل با حد پایین صفر است. بنابراین α های متقارن حول $\alpha = 0.8$ ، در حقیقت بیانگر حد پایین متقارن حول نقطه صفر هستند. با توجه به این مطلب می‌توان گفت که α های متقارن در حقیقت درجه آزادی مشابهی به شبکه می‌بخشند و از این جهت انتظار می‌رود رفتار یکسانی نیز داشته باشند. در شکل (۶)، منحنی تغییرات خطای متوسط EBP با $\alpha = 0.5$ و $\alpha = 1.1$ نشان داده شده است. شباهت دو منحنی می‌تواند بیانگر صحت مطلب فوق باشد.

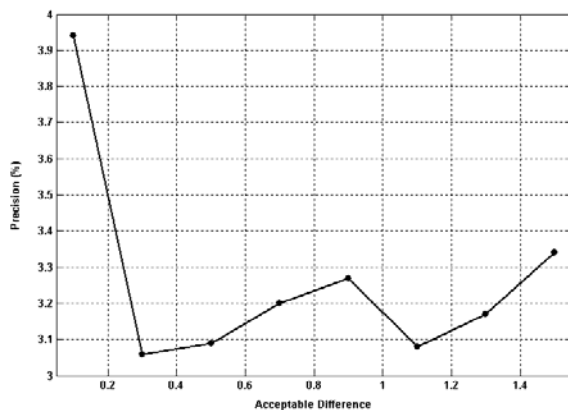
۵- نتیجه‌گیری

در این مقاله روشی جدید برای آموزش شبکه عصبی MLP در کاربردهای دسته‌بندی، با نام EBP ارائه شده است. روش پیشنهادی برای دسته‌بندی پیکسل‌های تصاویر ماهواره‌ای استفاده شده است. نتایج حاصل از EBP با

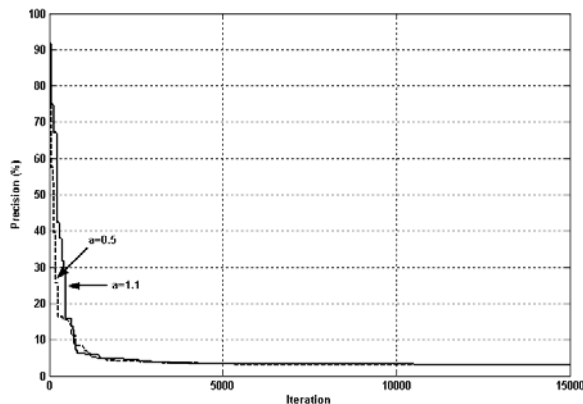
- [3] J.-C. Luo, C.-H. Zhou, Y. Leung, "A knowledge-integrated RBF network for remote sensing classification," 22nd Asian Conf. Remote Sensing, 2001.
- [4] C. Cortes and V. Vapnik, "Support vector networks," Machine Learning, 20:273-297, 1995.
- [5] B.Solaiman, M.C.Mouchot, and E.Maillard, "A hybrid algorithm, HLVQ, combining unsupervised and supervised learning approaches," Int. Conf. Neural Networks (ICNN'94), 1994.
- [6] J. D. Paola, R. A. Schowengerdt, "A detailed comparison of back-propagation neural network and maximum likelihood classifiers for urban land use classification," IEEE Trans. Geoscience and Remote Sensing, 33(4):981-996, 1995.
- [7] M. Lu, "Improved neural network modeling approach for engineering applications," 9th Int. Conf. Neural Information Processing, pp. 1810-1814, 2002.
- [8] S. Haykin, Neural networks: A comprehensive Foundation. Prentice-Hall, 1999.
- [10] J. Y. F. Yam and T. W. S. Chow, "Extended least squares based algorithm for training feedforward networks," IEEE Trans. Neural Networks, 8(3):806-810, 1997.
- [11] M. T. Hagan and M. B. Menhaj, "Training feedforward networks with the Marquardt algorithm," IEEE Trans. Neural Networks, 5(6):989-993, 1994.

جدول ۱: دقت نهایی الگوریتم‌های EBP, BP, ELS و ML به تفکیک کلاس.

	Class 1 (%)	Class 2 (%)	Class 3 (%)	Class 4 (%)	Class 5 (%)	Class 6 (%)	Class 7 (%)	Class 8 (%)	Total (%)
EBP($\alpha=0.1$)	1.46	1.79	8.62	2.87	0.0	12.12	9.00	4.87	3.94
EBP($\alpha=0.3$)	0.87	1.08	7.62	2.25	0.0	8.75	9.25	3.12	3.06
EBP($\alpha=0.5$)	0.62	1.29	6.25	2.25	0.0	10.25	9.37	3.37	3.09
EBP($\alpha=0.7$)	0.75	1.17	6.62	1.37	0.0	13.50	8.75	2.75	3.20
EBP($\alpha=0.9$)	0.62	1.29	7.87	1.87	0.12	11.50	10.25	3.00	3.27
EBP($\alpha=1.1$)	0.62	1.33	6.25	1.62	0.12	15.25	6.75	2.87	3.08
EBP($\alpha=1.3$)	0.67	1.25	6.00	1.75	0.0	14.25	7.87	2.50	3.17
EBP($\alpha=1.5$)	0.62	1.42	8.37	2.25	0.12	12.62	11.62	2.87	3.34
BP	0.67	1.42	7.50	2.25	0.0	13.75	10.50	3.12	3.35
ELS	1.00	1.42	9.5	3.62	0.0	12.75	12.62	6.12	4.21
ML	1.67	1.83	5.00	2.88	0.50	16.51	10.13	1.25	3.90



شکل (۵): منحنی دقت نهایی الگوریتم پیشنهادی بر حسب α .



شکل ۶: منحنی دقت نهایی الگوریتم پیشنهادی به ازای $\alpha = 0.5$ و $\alpha = 1.1$. شباهت دو منحنی می‌تواند بیانگر عملکرد یکسان شبکه به ازای α های متقارن باشد.