



ارزیابی قابلیت مدل شبکه عصبی مصنوعی LVQ4a2 در پیش بینی الگوی توزیع مکانی کادمیم در خاک

هادی قربانی^{۱*}- عباس روحانی^۲- ناصر حافظی مقدس^۳

تاریخ دریافت: ۱۳۹۰/۱۲/۲۴

تاریخ پذیرش: ۱۳۹۱/۱۰/۱۰

چکیده

این پژوهش به منظور پیش بینی کلاس بندی الگوی تغییرات مکانی عنصر کادمیم با استفاده از شبکه عصبی LVQ4a2 در سطح استان گلستان انجام شد. داده های مربوط به غلظت کادمیم از طریق نمونه برداری از سطح منطقه مورد مطالعه و اندازه گیری کادمیم کل خاک به دست آمد. برای ارزیابی قابلیت شبکه عصبی LVQ4a2 در پیش بینی توزیع مکانی عنصر کادمیم از مقایسه آماری پارامترهای مانند میانگین، واریانس، توزیع آماری رگرسیونی بین مقادیر پیش بینی شده مکانی توسط شبکه عصبی و مقادیر واقعی آنها و نیز معیار دقت کلاس بندی استفاده شد. نتایج حاصل نشان داد که در فاز آموزش و آزمایش تفاوت معنی دار در سطح ۵ درصد بین میانگین، واریانس و توزیع آماری مجموعه داده های پیش بینی شده مکانی عنصر کادمیم و مقادیر واقعی آنها مشاهده نشد. بنابراین شبکه عصبی مصنوعی به خوبی توانست مدل داده های مکانی عنصر کادمیم را بیاموزد. نتایج به دست آمده نشان داد که شبکه عصبی آموزش دیده، دارای قابلیت بالایی در پیش بینی مکانی غلظت کادمیم در نقاط نمونه برداری نشده با دقت تشخیص مساوی صفر درصد بود. نتایج این تحقیق نشان داد که شبکه عصبی مصنوعی می تواند بطور مناسبی در تشخیص سطوح مختلف غلظت کادمیم خاک بکار برده شود.

واژه های کلیدی: کلاس بندی، شبکه عصبی مصنوعی، کادمیم، خاک، استان گلستان

مقدمه

طبيعت دارای مقادير جزئي کادمیم می باشد، لیکن بخش بسیار عمده تجمع کادمیم در خاک مربوط فعالیت ها و منابع انسانی است (۲۸) به طوری که گفته می شود بیش از ۹۰ درصد کادمیم موجود در خاک های بسیاری از کشورهای غربی تنها ناشی از کاربرد کودهای فسفاته و رسوبات اتمسفری است (۹). از آنجا که ورود کادمیم به زنجیره غذایی انسان از طریق جذب و تجمع در بافت های گیاهی به سهولت امکان پذیر می باشد، لذا حتی مقادیر کم کادمیم در خاک نیز ممکن است در دراز مدت در سلامت محصولات کشاورزی و انسان ایجاد مشکل نماید. در همین راستا پژوهشگران در سال های اخیر ارزیابی خطر ناشی از آن معطوف داشته اند. امینی و همکاران (۱۱) با پنهنه بندی و ارزیابی خطرات ناشی از کادمیم و سرب در سلامت انسان در خاک های مناطق مرکزی ایران به این نتیجه رسیدند که خطرات ناشی از کادمیم در خاک های غرب منطقه مورد مطالعه بیشتر از سایر مناطق بوده و غلظت سرب در منطقه مورد مطالعه کمتر از حد استانه بروز خطر بوده است. در تحقیق دیگری امینی و همکاران

منابع مختلفی از مواد آلاینده در طبیعت وجود دارد. لیکن فلزات سنگین از جمله منابع مهم در این زمینه به شمار می آیند و از آنجا که اثرات مستقیم بر تولید محصولات کشاورزی و نیز زنجیره غذایی انسان دارد، توجه به آنها از اهمیت خاصی برخوردار است. کادمیم از جمله فلزات سنگینی است که به دلیل اهمیت زیست محیطی آن توجه بسیاری از محققین را در نقاط مختلف جهان و طی چندین دهه گذشته به خود جلب کرده است (۲۸). در سال های اخیر نیز محققین نشان داده اند که کادمیم می تواند علاوه بر تجمع پایدار در بافت های کلیوی که باعث رسوب پروتئین و صدمات جبران ناپذیر به کلیه ها می شود، افزایش خطر ابتلا به سرطان های پروستات را نیز به دنبال داشته باشد (۱۶ و ۲۷). هر چند سنگ ها و کانی های موجود در

۱- استادیاران گروه آب و خاک، دانشکده کشاورزی، دانشگاه صنعتی شاهرود
۲- نویسنده مسئول: Email: ghorbani1969@yahoo.com
۳- دانشیار دانشکده علوم، دانشگاه فردوسی مشهد

ترکیب شیمیایی بربوفیت‌های آبی و نوع بستر رودخانه (۳۲)، برآورد الگوی جذب فلزات سنگین در خاک‌های آلمان (۱۳)، توزیع عناصر رادیو اکتیو در خاک و ارتباط ویژگی‌های خاک با آنها (۱۷) و نیز ارزیابی کیفیت خاک از نظر ویژگی‌های بیولوژیکی (۲۶) به خوبی می‌تواند در کلاسه بندی، توزیع و تشریح بسیاری از پدیده‌های زیست محیطی در خاک مناسب بوده و به عنوان یک ابزار مناسب در این راستا مورد استفاده قرار گیرد. هدف این مطالعه، ارزیابی قابلیت مدل شبکه عصبی چندی ساز بردار یادگیر در پیش‌بینی، طبقه بندی و ترسیم نقشه‌های توزیع مکانی عنصر کادمیوم در نقاط نمونه برداری نشده بر اساس داده‌های حاصل از نقاط نمونه برداری شده در سطح منطقه مورد مطالعه و نیز پیدا کردن پارامترهای بهینه در چنین کاربردی می‌باشد.

مواد و روش‌ها

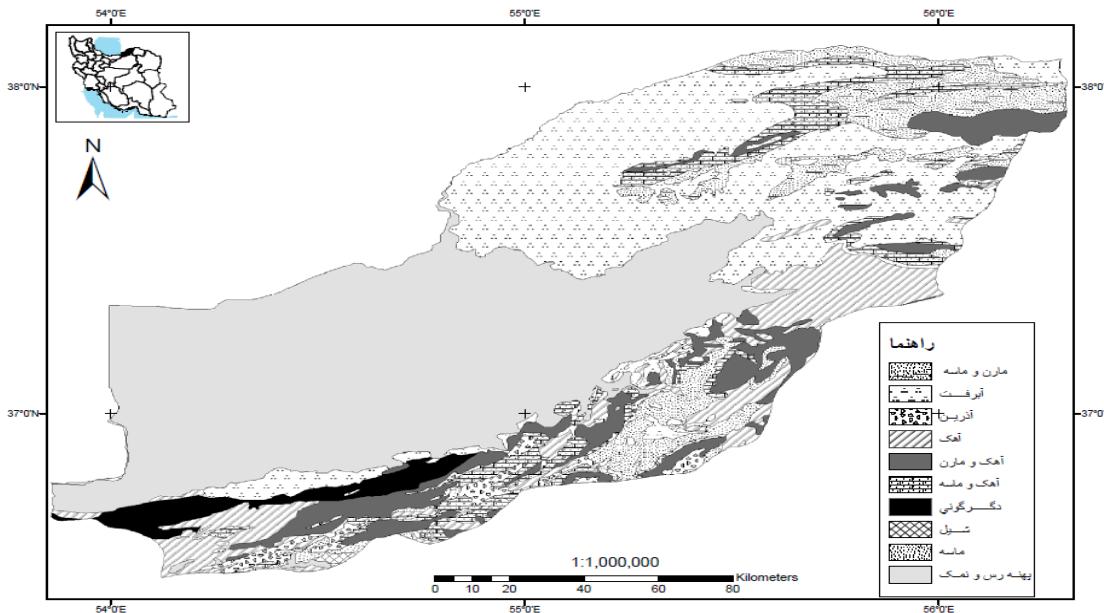
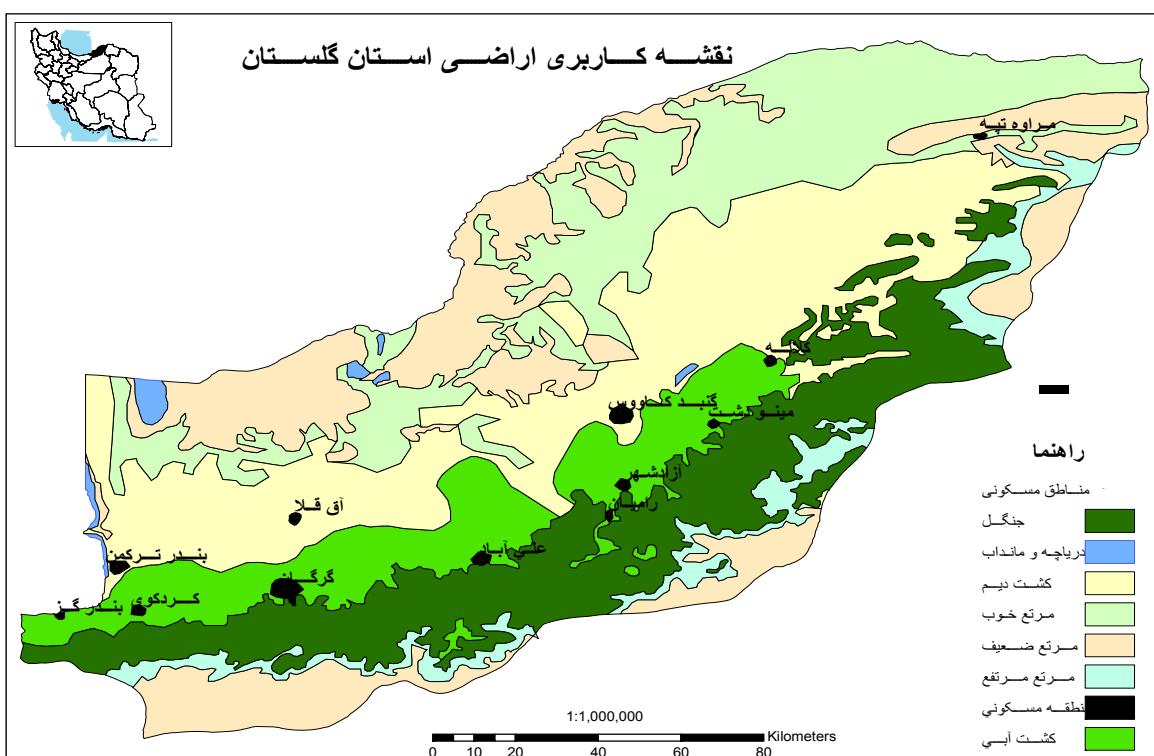
ویژگی‌های منطقه مورد مطالعه

محدوده مورد مطالعه، استان گلستان با مساحتی بالغ بر ۲۰۳۱۲ کیلومتر مربع، حدود ۱/۳ درصد از مساحت کل کشور را شامل می‌شود. این استان در مختصات جغرافیایی ۵۳ درجه و ۵۱ دقیقه تا ۵۶ درجه و ۱۴ دقیقه طول شرقی و ۳۶ درجه و ۴۴ دقیقه تا ۳۸ درجه و ۵ دقیقه عرض شمالی واقع شده است. استان گلستان از نظر تقسیم بندهای ایالات زمین ساختی، در پهنه زون‌های گرگان - رشت، البرز - آذربایجان و کپه‌داغ قرار دارد. بیشترین پهنه زون گرگان - رشت را جله گرگان پوشانیده است. در قسمت شرق، این زون‌ها با زون‌های بینالود و هزار مسجد کپه‌داغ هم مرز است که مرز آن با کپه‌داغ، پوشیده از بادرفت‌های خصیمی از جنس لس می‌باشد. وجود انواع نهشته‌ها به خصوص سنگ‌های رسوبی از دوره زمین‌شناسی پرکامبرین تا دوره کتونی این منطقه را از نظر ویژگی‌های زمین‌شناسی در دریف مناطق استثنایی قرار داده است. قدیمی‌ترین سنگ‌های استان شیسته‌های گرگان است (شکل ۱). از نظر ریخت شناسی حدود یک سوم استان کوهستانی و مابقی تپه ماهور و دشت می‌باشد. در جدول ۱ درصد واحدهای مختلف ریخت شناسی در استان آمده است. بخش عمده ای از دشت‌های استان زیر کشت آبی یا دیم می‌باشد. در شکل ۲ نقشه کاربری اراضی و پوشش گیاهی منطقه مورد مطالعه که با استفاده از داده‌های باند ۱ و ۲ سنجنده MODIS (با قدرت تفکیک مکانی ۲۵۰ متر) مربوط به ۲۰ دوره ۸ روزه در طول سال ۲۰۰۵ تهیه شده نشان داده شده است. ملاحظه می‌گردد که پوشش گیاهی استان با کاهش بارش از جنوب به سمت شمال به ترتیب از پهنه‌های جنگلی به مزارع آبی، مزارع دیم، مراتع خوب و مراتع ضعیف تغییر می‌کند.

(۱۲) با مطالعه الگوی توزیع کادمیوم قابل دسترس و ارتباط آن با ویژگی‌های خاک و نوع کاربری اراضی درخاکهای مناطق مرکزی و خشک ایران دریافتند که در ۹۰ درصد موارد غلظت کادمیوم بالاتر از استاندارد ۰/۰ میلی گرم بر کیلوگرم خاک (استاندارد مورد استفاده در کشور سوئیس) بود. همچنین نوع کاربری اراضی ارتباط معنی داری با غلظت کادمیوم کل در خاک داشت در حالی که با کادمیوم قابل استخراج با DTPA مرتبط نبود. از آنجا که توزیع مکانی آلینده‌ها از جمله کادمیوم در خاک به نوع و شدت منابع آلینده در آن منطقه مربوط است، اطلاع از الگوی توزیع مکانی کادمیوم در خاک می‌تواند تا حدود زیادی به مدیریت کشت و کار و عملیات اصلاحی و در نتیجه کاهش خطرات زیست محیطی کمک نماید. امروزه استفاده از تکنیک‌ها و ابزارهای پیشرفته نظیر شبکه‌های عصبی مصنوعی جهت شناسایی مناطق آلوده و نیز مطالعه الگوی توزیع مکانی آلینده‌ها در مناطق مختلف جهان گسترش قابل توجهی یافته است (۱۹، ۲۰، ۲۱، ۲۵ و ۲۹، ۳۵). کاظمی و حسینی (۲۲) با مقایسه روش‌های مختلف میان یابی به مطالعه غلظت برخی فلزات سنگین در رسوبات دریایی خزر پرداختند. همچنین برخی محققین دیگر از روش‌های شبکه‌های عصبی مصنوعی در انتخاب بهترین شاخص آلودگی در حاشیه جاده‌های پرترافیک استفاده کردند (۳۱). ارزیابی آلودگی ناشی از فلزات سنگین در رسوبات لاپروی شده با استفاده از روش‌های آزمایشگاهی عصاره گیری پی در پی و نیز شبکه عصبی مصنوعی توسط آریاس و همکاران (۱۵) نشان داد که روش‌های شبکه عصبی می‌تواند در این ارزیابی به طور موفقیت آمیز عمل نماید. همچنین در تحقیق دیگری نشان داد که شبکه عصبی مصنوعی بخوبی می‌تواند با تلفیق متغیرهای فیزیکی، شیمیایی و اکولوژیکی رسوبات در کلاس بنده رسوبات از جهت کیفیت آنها مفید باشد (۱۰). دیانی و همکاران (۲) جهت تخمین و پهنه‌بندی غلظت سرب، روی و کادمیوم در جنوب اصفهان از اطلاعات ماهواره لنده است و روش‌های رگرسیونی استفاده کردند. بررسی تغییرات مکانی سوری خاک در منطقه یزد و اردکان با استفاده از تصاویر ماهواره ای (۴)، مدل سازی شبکه عصبی مصنوعی برای پیش‌بینی مکانی سوری خاک (۵)، پیش‌بینی مکانی مقدار پروتئین دانه گندم با استفاده از مدل رقومی پستی و بلندی زمین در منطقه اردل استان چهارمحال و بختیاری و مقایسه روش‌های رگرسیونی و شبکه عصبی مصنوعی (۸) و نیز توسعه توابع انتقالی تخمین ظرفیت تبادل کاتیونی خاک با استفاده از رگرسیون و شبکه عصبی و اثر تفکیک داده‌ها بر دقت و صحت توابع (۷) از دیگر تحقیقات در استفاده از شبکه‌های عصبی در ایران می‌باشد. نتایج شبکه‌های عصبی مصنوعی در مدل سازی جذب سطحی سرب از محلول‌های آبی توسط کاج پسته (۳۶)، مدل سازی پالایش سرب از فاضلاب توسط روش‌های الکترودیالیز (۳۰)، کلاس بنده ارتباط بین

جدول ۱- درصد واحدهای مختلف ریخت شناسی در منطقه مورد مطالعه (۱)

واحد ریخت شناسی کوهستان	تیه ماهور	دشت رودخانه مانداب، باتلاق و دریاچه
۰/۶	۸/۴۸	۳۸/۸۳
	۲۰/۳۷	۳۱/۷
		درصد

**شکل ۱- نقشه زمین شناسی استان گلستان(برگرفته از نقشه های ۱:۲۵۰۰۰۰ و ۱:۱۰۰۰۰۰ سازمان زمین شناسی)****شکل ۲- نقشه کاربری اراضی و پوشش گیاهی استان گلستان(۱)**

نشود ادامه خواهد یافت. از بین شش گونه الگوریتم آموزش با نام‌های LVQ1، LVQ2.1، LVQ23، LVQ4 و LVQ4(24,23) OLVQ از LVQ4a2 به جهت بالا بودن کارآیی آن استفاده شد (۳۴ و ۳۳).

الگوریتم آموزش LVQ4a2

اگر فرض شود x^q بردار ورودی از مجموعه آموزش باشد: Q, \dots

$$\begin{aligned} & x^q \in Q, \quad q=1, \\ & w_m \in \mathcal{F}, \quad m=1, \dots, M \\ & \text{بردار کد باشد} \\ & \text{نزدیک ترین بردار کد به } x^q \text{ عبارت است از:} \\ & c = \arg \min_m (\|x^q - w_m\|) \end{aligned} \quad (1)$$

تنظیم: w_c

$$w_c(t+1) = w_c(t) + \alpha(n)s(n)[x^q - w_c(t)] \quad (2)$$

آنچه باید گیری در طول زمان آموزش از رابطه زیر به دست آمد.

$$a(n) = \begin{cases} k_1 & \text{if } q \leq q_c \\ k_2 \exp(-\frac{q-q_c}{r}) & \text{if } q > q_c \\ 0 & a(n) < 1 \end{cases} \quad (3)$$

اگر کلاس بندی صحیح باشد $d_c(n) = s(n)$ می‌باشد و در صورتی که کلاس بندی صحیح نباشد $-1 = s(n)$ خواهد شد. $d_c(n)$ عامل متعادل کننده است و از رابطه زیر به دست می‌آید:

$$d_c(n) = \frac{s}{p} \quad (4)$$

در این رابطه، به ترتیب (n) و $P_c(n)$ تعداد بردارهایی است که به درستی و اشتباه توسط بردار کد c کلاس بندی می‌شود. تا پایان تکرار اول مقدار $d_c(n)$ مشخص نخواهد شد. بنابراین از روش زیر برای محاسبه آن استفاده شد:

(5)

$$d_c(n) = \begin{cases} \frac{s}{\lambda} & \text{if } q = 1 \\ \frac{s}{d_c(n-1)} & \text{if } q \geq 2, P_c(n-1) \neq 0 \\ \lambda & \text{if } q \geq 2, P_c(n-1) = 0 \end{cases} \quad (2)$$

و $\lambda \geq 1$ می‌باشد. مقادیر توصیه شده $= 0.2$ و $= 10$ است.

بر اساس استانداردهای موجود غلظت تمام نمونه‌های برداشت شده کمتر از مقدار حد آستانه برای آلدگی بود. از آنجا که هدف از این تحقیق ارزیابی قابلیت پیش‌بینی و منطقه بندی شبکه عصبی در منطقه مورد مطالعه با غلظت‌های مشخص از عنصر کادمیم بود لذا چند مقدار غلظت فرضی جهت کلاس بندی در نظر گرفته شد.

روش نمونه برداری

جهت نمونه برداری از سطح استان ابتدا محدوده مورد مطالعه به شبکه‌های منظم مربع شکل با ابعاد $2/5$ کیلومتر و در مجموع ۱۷۰۰ شبکه تقسیم گردید. سپس با توجه به نقشه‌های خاک شناسی، زمین شناسی، کاربری اراضی، و نیز توزیع منابع آلینده شامل موقعیت معدن، صنایع، شبکه جاده‌های اصلی، محل‌های دفن زباله، زمین‌های کشاورزی آبی و دیم و غیره شبکه‌های با شرایط یکسان بطور ساده در هم ادغام گردید و نهایتاً از ۲۳۰ شبکه باقی مانده با توجه به منابع آلینده و پتانسیل آلدگی، نمونه‌های خاک سطحی از عمق ۰ تا ۳۰ سانتی متر تهیه گردید. در شکل ۳ موقعیت و تعداد نقاط نمونه گیری در شبکه‌های ۱۰ کیلومتری و براساس نوع آلدگی مورد انتظار در هر شبکه نشان داده است. با توجه به اطلاعات پایه و نقشه‌های اولیه موجود، ضمن تعیین انواع کاربری اراضی در منطقه، با توجه به نوع کاربری، محدوده مورد مطالعه به مناطقی با احتمال آلدگی کشاورزی، صنعتی و طبیعی تقسیم بندی گردید. بدینه است با توجه به کاربری اراضی و وضع موجود مناطق مختلف ممکن است دارای پتانسیل یک نوع آلدگی و یا بیشتر را توانما دارا باشند. در شکل ۳ در شبکه‌هایی که هر سه نوع آلدگی احتمالی وجود دارد (۱۶ نمونه)، در شبکه‌هایی با دو نوع آلدگی احتمالی چهار نمونه، در شبکه‌هایی با یک نوع آلدگی یک نمونه و در شبکه‌های بدون آلدگی مهم به ازاء هر چهار شبکه یک نمونه برداشت شده است.

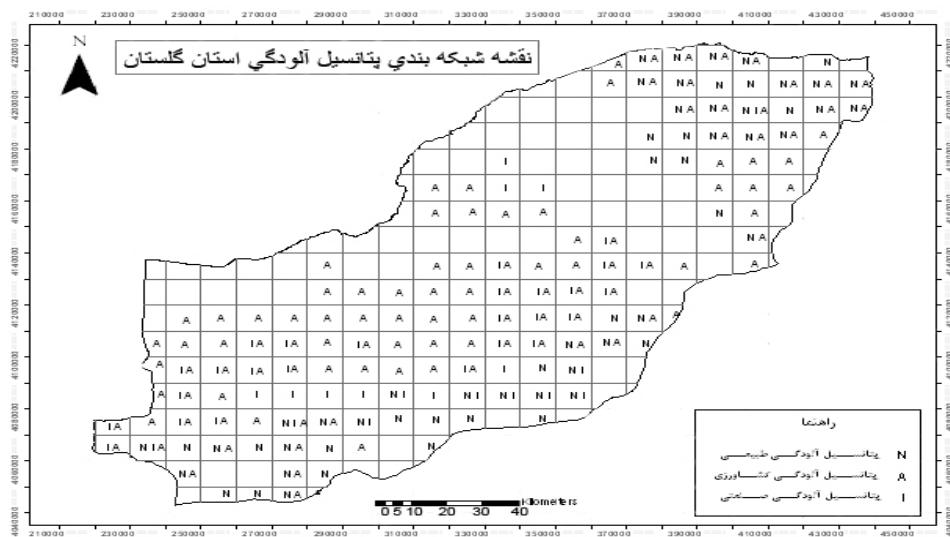
نمونه‌ها پس از خشک شدن در هوای آزاد (هوای خشک)، کوبیده شده و از الک ۲ میلی متري عبور داده شدند. غلظت کل کادمیم موجود در نمونه‌های خاک پس از آماده سازی، در آزمایشگاه LabWest (Aqua Regia) (ICP) مورد اندازه گیری قرار گرفت. داده‌های کادمیم بدست آمده از منطقه مورد مطالعه در فرآیند تجزیه و تحلیل شبکه عصبی مصنوعی مورد استفاده قرار گرفت.

شبکه عصبی LVQ4a2

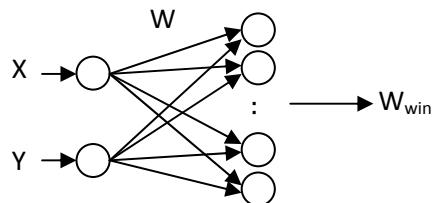
شبکه عصبی LVQ4a2 یکی از روش‌های طبقه بندی و خوش بندی است. نحوه ای کار آن بر اساس تقسیم فضای ورودی R^n به تعدادی منطقه‌های متمایز از هم با نام ناحیه‌های تصمیم گیری^۱ و اختصاص یک بردار کد به هر یک از آنها است. کلاس بندی بر حسب نزدیکی بردار ورودی x به بردارهای کد انجام می‌شود. بردار ورودی متعلق به کلاس نزدیک ترین بردار کد گذاری خواهد شد. الگوریتم تا وقتی که ناحیه‌های تصمیم گیری و مرکزهای آنها ثابت

1 - Decision regions (Voronoi cells)

2 - Codebook (Voronoi) vector



شکل ۳- نقشه شبکه بندی پتانسیل آلودگی و راهنمای تهیه نمونه‌های خاک در منطقه مورد مطالعه



شکل ۴- ساختار شبکه LVQ استفاده شده برای پیش‌بینی مکانی عنصر کادمیم. (X,Y) مختصات نقاط نمونه برداری شده، W ماتریس بردارهای کد و W_{win} برچسب بردار کد برنده شده

و کولموگرف- اسمیرنوف برای مقایسه بین میانگین، واریانس و توزیع آماری مقادیر پیش‌بینی شده و مقادیر واقعی آنها و نیز رابطه خطی رگرسیونی بین آنها انجام شد (۳ و ۶). برنامه نویسی این مدل در محیط نرم افزار Matlab نسخه ۷ انجام شد (۱۴).

نتایج و بحث

خطای تشخیص شبکه عصبی LVQ4a2 در دو مرحله آموزش و آزمایش در هر مورد طبقه بندی منطقه در جدول ۲ نشان داده شده است. خطای تشخیص، یعنی تعداد الگوهایی که به غلط توسط شبکه عصبی کلاس بندی شده به تعداد کل الگوهای موجود، در دو مرحله کاری شبکه در این جدول مشاهده می‌شود. به راحتی می‌توان براساس مقدار پایین بودن خطای تشخیص در فاز آموزش به شبکه عصبی برای کلاس بندی سطح منطقه اطمینان حاصل کرد و می‌توان بیان کرد که شبکه عصبی آموزش دیده به طور کامل الگوهای موجود را یاد گرفته است. بنابراین وزن‌های شبکه (بردارهای

هدف در استفاده از شبکه LVQ4a2 کلاس بندی منطقه به دو کلاس است. کلاس اول نقاطی از منطقه که دارای غلظت‌های عنصر کادمیم کمتر از مقادیر میانگین کل غلظت‌های نمونه‌های برداشت شده، نصف میانگین و سه چهارم میانگین است و نقاط دیگری که بیشتر یا مساوی مقادیر میانگین، نصف میانگین و سه چهارم میانگین می‌باشند. ساختار شبکه عصبی LVQ4a2 استفاده شده به منظور کلاس بندی منطقه بر مبنی غلظت آلودگی در شکل ۴ نشان داده شده است.

ورودی‌های شبکه مختصات مکانی می‌باشند و خروجی شبکه بردار کدی است که نزدیکترین فاصله اقیദسی را با ورودی دارد. برچسب بردار کد تعیین کننده کلاس بردار ورودی است. به طور کاملاً تصادفی، از بین 230° داده موجود از نقاط نمونه برداری شده از سطح منطقه 80° درصد از کل داده‌ها برای آموزش شبکه و 20° درصد باقی مانده برای آزمایش شبکه بمنظور کلاس بندی انتخاب شد. برای ارزیابی قابلیت پیش‌بینی شبکه عصبی به ترتیب از آزمون F،

پراکنش عنصر کادمیم می‌باشد. وجود $p=1$ در مورد توزیع آماری بین مقادیر پیش‌بینی شده و مقادیر واقعی عنصر کادمیم در سطح منطقه مورد مطالعه به ویژه در فاز آزمایش شبکه، نشان از دقت بالا و نیز قابلیت خوب شبکه عصبی مصنوعی LVQ4a2 به منظور کلاس بندی سطح منطقه از نظر مقادیر مشخص شده می‌باشد. کاظمی و حسینی (۲۲) نیز با مقایسه چهار مدل درونیابی داده‌ها دریافتند که الگوریتم ژنتیکی مبتنی بر شبکه عصبی مصنوعی توانست بهترین مدل در تخمین و پیش‌بینی غلظت برخی فلزات سنگین در رسوبات دریای خزر در برخی کشورهای حاشیه این دریا باشد.

ضرایب تبیین و همچنین رابطه خطی رگرسیونی بین مقادیر واقعی هر کلاس در مقابل مقادیر پیش‌بینی شده توسط شبکه عصبی LVQ4a2 در جدول ۴ نشان داده شده است.

بهترین نتایج براساس این دو معیار وقته به دست می‌آید که معادله خطی بین غلظت واقعی و غلظت پیش‌بینی شده توسط شبکه عصبی علاوه بر داشتن ضرایب تبیین بالا، دارای کمترین عرض از مبدأ (نزدیک به صفر) و شبیه نزدیک به یک باشد ($p_{v}=1/000$, $d_{v+0}=1/000$).

کاملاً واضح است که ضرایب تبیین بین داده‌های واقعی و پیش‌بینی شده بسیار بالا می‌باشند ($R^2 > 0.8$). همچنین معادله خطی رگرسیونی بین آنها دارای شبیه های نزدیک به واحد و عرض از مبدأهای کوچکی هستند، بنابراین می‌توان به چنین شبکه‌های اطمینان کرد.

کد) مقدار بهینه داشتند. در مرحله بعدی، ویژگی تعمیم پذیری و سایستگی شبکه عصبی آموزش دیده ارزیابی شد. بنابراین شبکه با استفاده از مجموعه داده‌های غیر از مجموعه داده‌های آموزش (مجموعه داده‌های تست) ارزیابی شد. خطای تشخیص در فاز آزمایش صفر بود که این نشان از ویژگی تعمیم پذیر شبکه عصبی LVQ4a2 برای سایر نقاط دارد. بنابراین می‌توان به شبکه عصبی آموزش دیده اطمینان داشت.

برای اثبات ادعای ما در مورد قابل اطمینان بودن شبکه‌های عصبی آموزش دیده برای پیش‌بینی الگوی پراکنش عنصر کادمیم، داده‌های واقعی و داده‌های پیش‌بینی شده توسط شبکه از نظر آماری با هم مقایسه شدند. در اینجا فرضیه صفر برتساوی میانگین، واریانس و توزیع آماری هر دو سری داده دلالت دارد. هر فرضیه در سطح احتمال ۹۵ درصد به کمک پارامتر p آزمون شد. بنابراین اگر محاسبه شده برای هر مقایسه بیشتر از 0.05 باشد، فرضیه صفر را نمی‌توان رد کرد. به ترتیب برای مقایسه میانگین، واریانس و توزیع آماری از آزمون t و کولموگرف- اسمیرنو استفاده شد. مقادیر محاسبه شده برای هر مورد کلاس بندی در جدول ۳ نشان داده شده است.

این نتایج نشان می‌دهند که میانگین، واریانس و توزیع آماری مقادیر واقعی و پیش‌بینی شده در هر مورد طبقه بندی، اختلاف معنی داری با هم نداشتند (در تمام موارد $0.05 < p$). بنابراین شبکه عصبی مصنوعی LVQ4a2 دارای قابلیت بالایی برای پیش‌بینی الگوی

جدول ۲- خطای تشخیص شبکه عصبی LVQ4a2 در دو فاز آموزش و آزمایش

هدف شبکه برای طبقه بندی	تعداد عضوهای کلاس یک	تعداد عضوهای کلاس دو	خطای تشخیص در مرحله آزمایش(%)	خطای تشخیص در مرحله آموزش(%)	آزمایش (%)
$D_{\text{mean}} = 0.064 \text{ ppm}$	۷۷	۱۵۱	۱/۳۹	۱/۰۰	۰/۰۰
$D_{\text{mean}} = 0.128 \text{ ppm}$	۱۱۹	۱۰۹	۲/۷۶	۰/۰۰	۰/۰۰
$D_{\text{mean}} = 0.192 \text{ ppm}$	۱۸۱	۴۷	۰/۹۲		

توضیحات: D_{mean} متوسط غلظت نمونه‌ها، $D_{\text{mean}}/2$ نصف متوسط غلظت نمونه‌ها، کلاس اول و کلاس دوم به ترتیب، اشاره به آن دسته از غلظت‌های دارد که کوچکتر از حد بحرانی فرضی و بزرگتر و مساوی حد بحرانی فرضی می‌باشد.

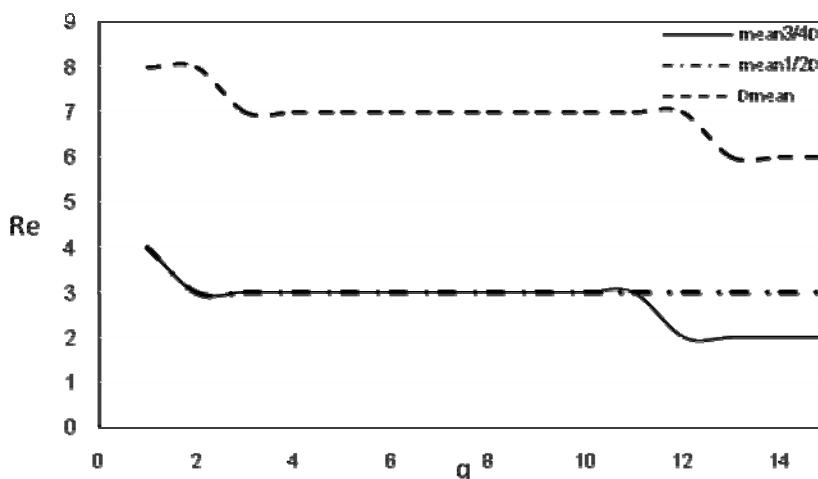
جدول ۳- مقایسات آماری مقادیر واقعی و پیش‌بینی شده

هدف شبکه برای طبقه بندی مرحله به کارگیری شبکه	مقایسه یانگین	مقایسه واریانس	مقایسه توزیع	نوع تحلیل آماری
D_{mean}				فال آموزش
				فال آزمایش
D_{mean}				فال آموزش
				فال آزمایش
D_{mean}				فال آموزش
D_{mean}				فال آزمایش

جدول ۴- رابطه خطی رگرسیونی و ضریب تبیین بین dv و pv

R^2	هدف شبکه برای طبقه بندی مرحله بکارگیری شبکه رابطه خطی رگرسیونی	$pv=1/... dv+.../...$	فاز آموزش	$D_{1/2\text{mean}}$
.۹۹۹			فاز آزمایش	
.۹۹۹			فاز آموزش	D_{mean}
.۸۷۸		$pv=-.۹۳۳ dv+/.۰۶$	فاز آموزش	
.۹۹۹		$pv=1/... dv+.../...$	فاز آزمایش	
.۹۷۳		$pv=-.۹۹۴ dv+/.۱۱$	فاز آموزش	$D_{4\text{mean}}$
.۹۹۹		$pv=1/... dv+.../...$	فاز آزمایش	

توضیحات: dv : مقادیر واقعی داده ها pv : مقادیر پیش بینی شده توسط مدل



شکل ۵- نمودار همگرایی شبکه عصبی در طول مدت آموزش

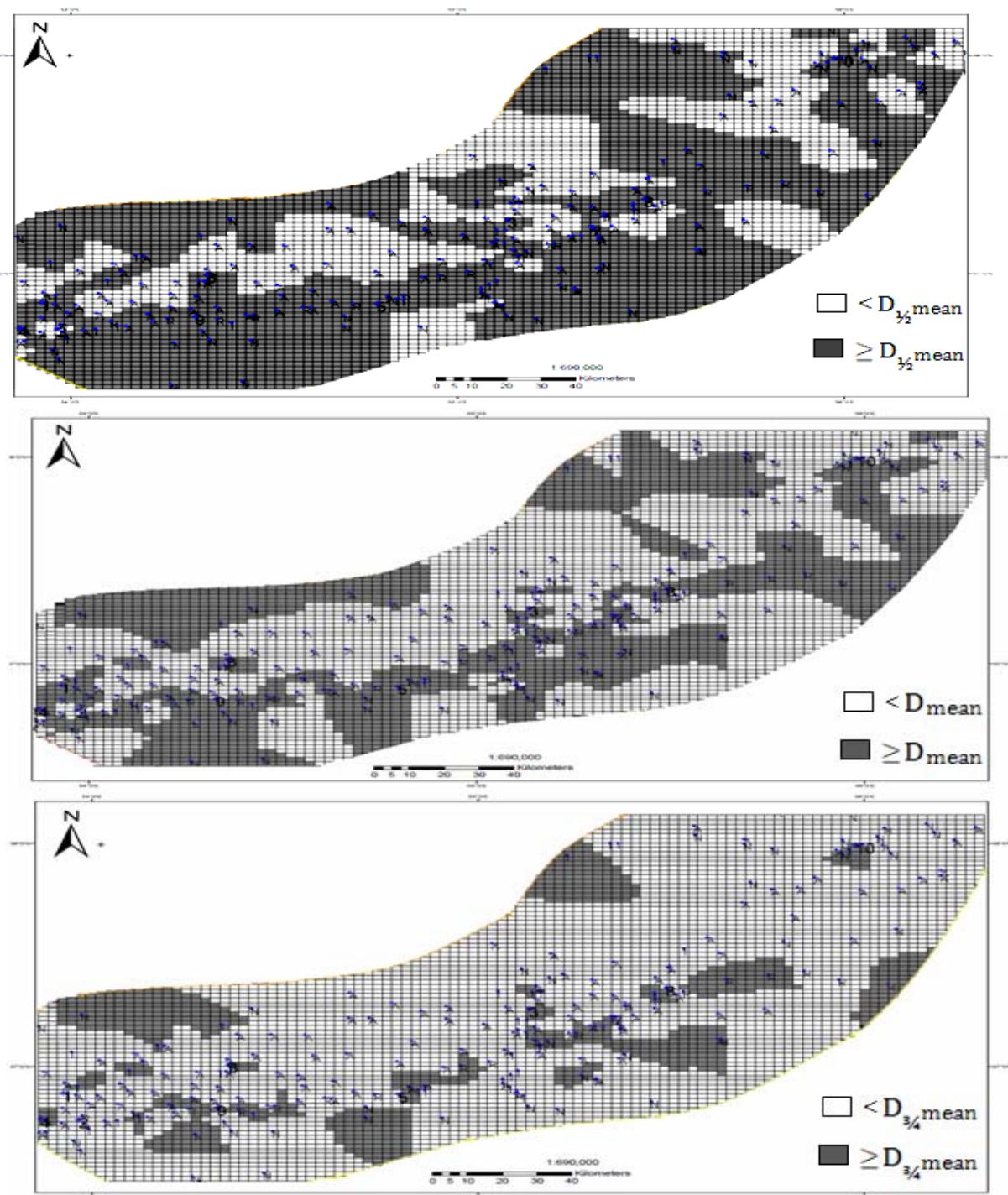
کادمیم را یاد بگیرد و آنها را پیش بینی کند و همچنین بر اساس نتایج مندرج در جدول های ۲، ۳ و ۴ از قابلیت تعیین پذیر بالایی نیز برخوردار است. قابلیت مناسب شبکه عصبی به عنوان ابزاری در طبقه بندی ارتباط بین خواص شیمیایی برخی گیاهان آبزی و ویژگی های بستر جریان های آبی (۳۲) و نیز در انتخاب بهترین نشانگر آلودگی ناشی از ترافیک جاده ای توسط سامانکا سایمر من و همکاران (۳۱) گزارش شده است.

نقشه توزیع مکانی کادمیم در منطقه مورد مطالعه
نقشه پراکنش مکانی عنصر کادمیم در منطقه مورد مطالعه که توسط شبکه عصبی LVQ4a2 تفکیک و ترسیم شده در شکل ۶ نشان داده شده است. در این شکل منطقه به ترتیب به سه کلاس با غلظت های کادمیم کمتر از مقدار میانگین (D_{mean})، نصف مقدار میانگین ($D_{1/2\text{mean}}$) و سه چهارم مقدار میانگین ($D_{3/4\text{mean}}$) و بیشتر یا مساوی مقادیر ذکر شده می تواند یک حد آستانه فرضی برای عنصر کادمیم باشد تا به کمک آن قابلیت شبکه عصبی مورد آزمایش قرار گیرد.

این نتایج نشان می دهند که قابلیت تعیین پذیری شبکه عصبی آموزش دیده بسیار بالا می باشد، زیرا فاز آزمایش در مقایسه با فاز آموزش دارای ضریب تبیین بالاتر و نیز معادله خطی رگرسیونی بین مقادیر واقعی و پیش بینی شده از شبکه نزدیکتر به واحد و عرض از مبدأ نزدیکتر به صفر برخوردار است.

نمودار همگرایی شبکه عصبی LVQ4a2 در مدت زمان آموزش برای هر مورد کلاس بندی در شکل ۵ نشان داده شده است. خطای تشخیص یا کلاس بندی برای هر مورد از تعداد ۴، ۸ و ۱۶ نقطه ($Re=4$) برای کل مجموعه آموزش، شروع و به تدریج کم تر می شود و بعد از ۱۳ تکرار ($q=13$)، خطای ثابت باقی می ماند زیرا مرکزها و ناحیه های تشخیص هر بردار کد بد邦ون تغییر می مانند. افزایش تعداد تکرارهای الگوریتم نیز موجب حرکت بردارهای کد به موقعیت دیگری نخواهد شد. لذا ادامه آموزش بی فایده خواهد بود و بنابراین آموزش شبکه باید متوقف شود. معیار توقف الگوریتم بر اساس همین فرض بنا نهاده شد. (q: تعداد تکرارهای آموزش شبکه و Re : خطای کلاس بندی در مرحله آموزش بر حسب تعداد)

نتایج حاصل از شکل ۵ نشان می دهند که شبکه عصبی LVQ4a2 در حداقل زمان ممکن می تواند الگوی پراکنش عنصر



شکل ۶- نقشه توزیع مکانی و تفکیک شده عنصر کادمیم براساس مقادیر آستانه فرضی توسط مدل شبکه عصبی LVQ4a2

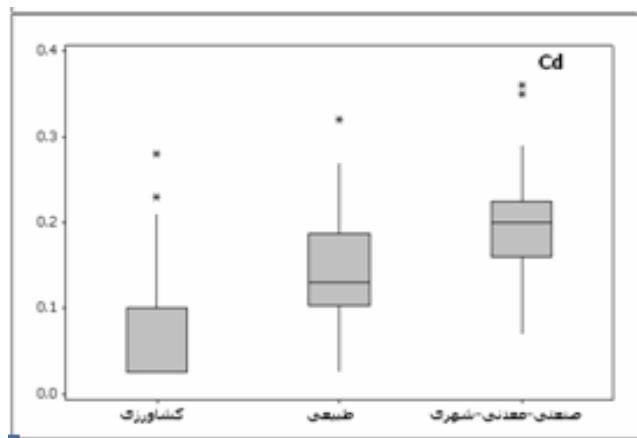
آلوده به عنصر کادمیم و نیز تفکیک آنها در صورت وجود آلودگی باشد. آریاس و همکاران (۱۵) نیز دریافتند که داده‌های آماری حاصل از شبکه عصبی مصنوعی ثابت کرد که مقادیر تخمینی فلزات سنگین نیکل، سرب و روی در رسوبات رودخانه دارای الگوی پراکنش بسیار

با توجه به آنکه پیشتر با مقایسات آماری مشخص که، خطای تشخیص و معادله رگرسیونی بین مقادیر واقعی و پیش‌بینی شده تفاوت معنی داری نداشتند، لذا می‌توان بیان کرد شبکه عصبی مصنوعی LVQ4a2 می‌تواند یک ابزار مفید جهت تشخیص مناطق

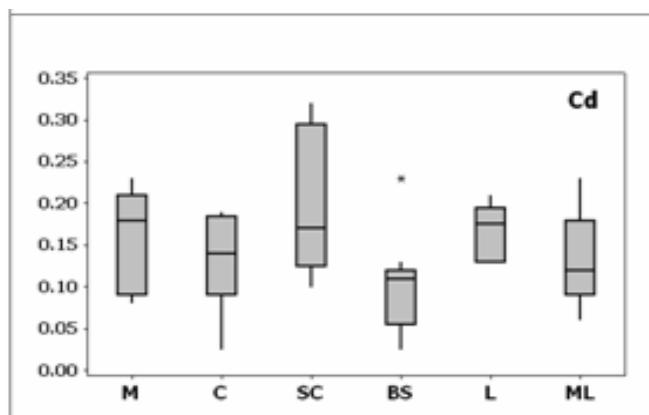
برخی دلایل تغییرات غلظت کادمیم در منطقه مورد مطالعه در شکل ۷ دیاگرام جعبه ای از تغییرات کادمیم به تفکیک در نمونه های مختلف برداشت شده از منطقه نشان داده شده است. ملاحظه می گردد که بیشترین غلظت کادمیم مربوط به نمونه های صنعتی، معدنی و شهری است و در مرحله بعد منشاء طبیعی غلظت های بالاتری را نشان می دهد.

همچنین در شکل ۸ نیز مشاهده می گردد که در بین نمونه های طبیعی بیشترین غلظت کادمیم اندازه گیری شده مربوط به سنگ های مافیک و لس ها می باشد و البته در سنگ های شیل و ماسه سنگی زغال دار تغییرات این عنصر بیشتر از سایر گروه ها است. نتایج آنالیز فاکتورها نیز نشان می دهد که عنصر کادمیم تنها همراه با سرب در یک گروه قرار می گیرند و منشاء انسان زاد این عنصر را تائید می کند. همچنین نتایج حاصل از نقشه پهنه بندی تغییرات کادمیم که به روش شبکه مصنوعی تهیه گردید (شکل ۶) و نیز نقشه پتانسیل آلودگی استان (شکل ۳).

مشابهی با مقادیر واقعی در منابع ورودی به آب است. کاربرد عملی نقشه های تهیه شده با شبکه عصبی LVQ4a2 ابزارهای موثر و دقیقی خواهد بود که به ما جهت پیش بینی آلودگی مناطق آلوده، پیش بینی روند آلودگی مکانی منطقه و نیز ردیابی منابع آلودگی در منطقه کمک خواهد کرد. آواراز گورا و همکاران (۱۰) نیز در تحقیق خود نتیجه گرفتند که شبکه عصبی مصنوعی SOM به خوبی قادر است تا با تلفیق داده های فیزیکی، شیمیایی و اکولوژیکی موجود اقدام به کلاس بندی رسوبات مختلف با کیفیت مشابه در منطقه مورد مطالعه نماید. آنان نتیجه گرفتند که شبکه عصبی مصنوعی مورد استفاده می تواند به عنوان یک ابزار قوی و قابل اعتماد در مطالعات مشابه مورد استفاده قرار گیرد. شکل ۶ نقشه های حاصل از تفکیک منطقه مورد مطالعه را که بر اساس حد آستانه D_{mean} و $D_{\% \text{ mean}}$ در تمام نقاط منطقه به کمک شبکه عصبی LVQ4a2 آموزش دیده ترسیم شده است، نشان می دهد.



شکل ۷- تغییرات غلظت کل کادمیم خاک (mg kg^{-1}) با توجه به کاربری اراضی در منطقه مورد مطالعه



شکل ۸- تغییرات غلظت کل کادمیم خاک (mg kg^{-1}) در خاک های ناشی از واحد سنگی مختلف (سنگ های مافیک (M)، سنگ های کربناته (C)، تناوب شیل و ماسه سنگ های زغالدار (SC)، شیل های سیاه (BS)، لس (L) و تالاب ها و گلفشان ها (ML))

گیاهان بر اساس قدرت آنها در جذب کادمیم و توانایی تجمع پذیری آن در بافت‌های گیاهی در مناطق مختلف با توجه به شدت و پراکنش کادمیم بنوان یک فلز آلاینده و خطرناک در خاک، تعیین و بطور قابل توجهی متفاوت خواهد بود. بر اساس نتایج حاصل از این تحقیق، شبکه‌های عصبی مصنوعی و بطور ویژه شبکه عصبی مصنوعی LVQ4a2 نشان داد که از قابلیت بالایی در تعمیم داده‌های آزمایش به سایر مناطق بوده و با دقت بالایی قادر به ترسیم نقشه‌های توزیع مکانی فلز سنگین کادمیم با توجه به حدود بحرانی و آستانه‌های در نظر گرفته شده در منطقه مورد مطالعه است. لذا شبکه عصبی مصنوعی مذکور می‌تواند در سایر مناطق و مطالعات مشابه به عنوان ابزار مفید و سریع بکار گرفته شده و با مدیریت مبتنی بر اطلاعات مکانی، در کاهش هزینه‌های ناشی از حضور کادمیم در خاک به طور موثری مفید واقع گردد.

سپاسگزاری

مولفین مقاله از سازمان حفاظت محیط زیست کشور و دانشگاه صنعتی شاهرود به جهت تامین منابع مالی و پشتیبانی این تحقیق تقدير و تشکر می‌نمایند.

که بر اساس نقشه جاده‌های اصلی استان، موقعیت معادن، واحدهای صنعتی و محلهای دفن زباله و مطالعات تکمیلی تهیه شده است، نشان داد که همخوانی مناسبی بین غلظت کادمیم و منابع انسانی بخصوص جاده‌ها وجود دارد (۱) که خود بیانگر تأثیر انسان بر الگوی توزیع مکانی کادمیم در منطقه مورد مطالعه می‌باشد.

نتیجه گیری

از آنجا که انتقال فلزات سنگین از جمله کادمیم از خاک به زنجیره غذایی انسان مشکلات بهداشتی و سلامتی قابل توجهی را برای انسان به دنبال خواهد داشت، مطالعه توزیع مکانی آن در خاک می‌تواند از توجه ویژه ای برخوردار باشد. اطلاع از توزیع مکانی غلظت آلاینده‌ها در مناطق مختلف می‌تواند ضمن پایش و کترول آنها در منطقه به مدیریت خاک کمک قابل توجهی نماید. از آنجا که هرگونه عملیات اصلاحی و یا مدیریت کشت و کار مبتنی بر اطلاعات و داده‌های خاک در مناطق مختلف می‌باشد لذا استفاده از ابزارهای مناسب و سریع در پردازش داده‌ها و تعمیم آنها به نقاط مجھول و ترسیم نقشه‌های پهنۀ بندی می‌تواند بسیار سودمند باشد. بطور مثال در منطقه مورد مطالعه با توجه به غلظت و تجمع کادمیم در نقاط مختلف، مدیریت خاک نیز متفاوت خواهد بود. بدینهی است عملیات و اصلاحات فیزیکی، شیمیایی و یا بیولوژیکی در خاک و یا انتخاب نوع

منابع

- حافظی مقدس ن، قربانی ۵، دهر ازما ب، رمضانی اومالی ر، الله فردوست ف، و قاسمی ح. ۱۳۸۹. مطالعات تهیه اطلس آلودگی خاک استان گلستان. گزارش نهایی، سازمان حفاظت محیط زیست ایران.
- دیانی م، نادری م، و محمدی ج. ۱۳۸۹. پهنۀ بندی غلظت سرب، روی و کادمیم در خاک با استفاده از داده‌های ماهواره لند ست ای ام در جنوب شهرستان اصفهان، مجله آب و خاک، سال بیست و چهارم، شماره ۲.
- روحانی ع. و مکاریان ح. ۱۳۹۰. تهیه نقشه‌های مدیریتی علف هرز با استفاده از شبکه‌های عصبی مصنوعی با هدف کاربرد در کشاورزی دقیق. نشریه ماشین‌های کشاورزی. شماره ۲، ۷۴-۸۳.
- روستایی صدرآبادی ف.س، فیض نیا ش، ایوبی س، علوی پناه ک. و نقیبی ح. ۱۳۸۷. بررسی تغییرات مکانی شوری خاک در منطقه یزد - اردکان با استفاده از تصاویر ماهواره ETMT. مجموعه مقالات یازدهمین کنگره علوم خاک ایران.
- روستایی صدرآبادی ف.س، فیض نیا ش، ایوبی س، علوی پناه ک. و نوروزی م. ۱۳۸۷. استفاده از مدل سازی شبکه عصبی مصنوعی برای پیش‌بینی مکانی شوری خاک در دشت یزد و اردکان با استفاده از داده‌های سنجنده ASTER. مجموعه مقالات یازدهمین کنگره علوم خاک ایران.
- مکاریان ح. و روحانی ع. ۱۳۸۹. پیش‌بینی الگوی توزیع جمعیت علف هرز جوموشی (*Hordeum glaucum* steud.) با استفاده از شبکه عصبی مصنوعی. نشریه حفاظت گیاهان (علوم و صنایع کشاورزی). جلد ۲۶، شماره ۴، ص ۴۷۱-۴۸۰.
- مهاجر ر، صالحی م.ح. و بیگی ح. ۱۳۸۸. توسعه توابع انتقالی تخمین ظرفیت تبادل کاتیونی خاک با استفاده از رگرسیون و شبکه عصبی و اثر تفکیک داده‌ها بر دقت و صحت توابع، مجله علوم و فنون کشاورزی و منابع طبیعی، دانشگاه صنعتی اصفهان، سال سیزدهم، شماره ۴۹، ۸۳ تا ۹۸.
- نوروزی م، جلالیان ا، ایوبی ش. و دهقانی ا.ا. ۱۳۸۷. پیش‌بینی مکانی مقدار پروتئین دانه گندم با استفاده از مدل رقومی پستی و بلندی زمین در

منطقه اردل استان چهارمحال و بختیاری (مقایسه روشها رگرسیونی و شبکه عصبی مصنوعی). مجموعه مقالات یازدهمین کنگره علوم خاک ایران.

- 9- Alloway B.J. 1995. Cadmium. In: Heavy metals in soils (pp. 122–151). Chapman and Hall, London.
- 10- Alvarez-Guerra M., González-Piñuela C., Andrés A., Galán B., and Viguri J.R. 2008. Assessment of Self-Organizing Map artificial neural networks for the classification of sediment quality. *Environment International* 34:782–790.
- 11- Amini M., Afyuni M., Khademi H., Abbaspour K.C., and Schulin R. 2005a. Mapping risk of cadmium and lead contamination to human health in soils of central Iran. *Science of The Total Environment*, 347(1-3): 64-77.
- 12- Amini M., Khademi H., Afyuni M., and Abbaspour K.C. 2005b. Variability of available cadmium in relation to soil properties and landuse in an arid region in central Iran. *Water, Air and Soil Pollution*, 162(1-4).
- 13- Anagu I., Ingwersen J., Utermann J., and Streck T. 2009. Estimation of heavy metal sorption in German soils using artificial neural networks. *Geoderma* 152: 104–112.
- 14- Anonymous. 2007. Matlab. The Language of Technical Computing. Version 7.4 . The Mathworks of Natick, Massachusetts, USA.
- 15- Arias R., Barona A., Ibarra-Berastegi G., Aranguiz I., and Elías A. 2008. Assessment of metal contamination in dredged sediments using fractionation and Self-Organizing Maps. *Journal of Hazardous Materials* 151: 78–85.
- 16- Benbrahim-Tallaa L., Liu J., Webber M.M., and Waalkes M.P. 2007. Estrogen signaling and disruption of androgen metabolism in acquired androgen-independence during cadmium carcinogenesis in human prostate epithelial cells. *The Prostate*, 67: 135–145.
- 17- Dragovic S., Mihailovic N., and Gajic B. 2008. Heavy metals in soils: Distribution, relationship with soil characteristics and radionuclides and multivariate assessment of contamination sources. *Chemosphere* 72: 491–495.
- 18- Ghorbani H. 2003. Sludge Application and Salinity on the Bioavailability of Cadmium in Soil. PhD Thesis, The University of Reading, United Kingdom.
- 19- Gómez-Carracedo M.P., Andrade J.M., Carrera G.V.S.M., Aires-de-Sousa J., Carlosena A., and Prada D. 2010. Combining Kohonen neural networks and variable selection by classification trees to cluster road soil samples. *Chemometrics and Intelligent Laboratory Systems* 102 : 20–34.
- 20- Gruszczynski S. 2007. The problem of the estimation of the industrial soil pollution extent. *Polish Journal of Soil Science* Vol. XL/1.
- 21- Kanevski M.F. 1999. Spatial Predictions of Soil Contamination Using General Regression Neural Networks. *Int. Journal of Systems Research and Information Systems*, 8(4): 241-256.
- 22- Kazemi S.M., and Hosseini S.M. 2011. Comparison of spatial interpolation methods for estimating heavy metals in sediments of Caspian Sea, *Expert Systems with Applications* 38:1632-1649.
- 23- Kohonen T. 1990 The Self-Organizing Map. *Proceedings of the IEEE*, 78:1464-1480.
- 24- Kohonen T. 1995 Self-Organizing Maps. Springer, Berlin.
- 25- Lentzsch P., Wieland R., and Wirth S. 2005. Application of multiple regression and neural network approaches for landscape-scale assessment of soil microbial biomass. *Soil Biology & Biochemistry* 37: 1577–1580.
- 26- Mele Pauline M., and Crowley D.E. 2008. Application of self-organizing maps for assessing soil biological quality. *Agriculture, Ecosystems and Environment* 126:139–152.
- 27- Miller W.R. 1996. Estrogen and breast cancer. Chapman and Hall, London.
- 28- Pan J., Plant J.A., Voulvoulis N., Oates C.J., and Ihlenfeld C. 2010. Cadmium levels in Europe: implications for human health. *Environ Geochem Health* 32:1–12.
- 29- Recknagel F., Bobbin J., Whigham P., and Wilson H. 2002. Comparative application of artificial neural networks and genetic algorithms for multivariate time-series modeling of algal blooms in freshwater lakes. *Journal of Hydroinformatics* : 04.2.2002.
- 30- Sadrzadeh M., Mohammadi T., Ivakpour J., and Kasiri N. 2006. Neural network modeling of Pb²⁺ removal from wastewater using electrodialysis. *Analytica Chimica Acta* 560 :172–183.
- 31- Samecka-Cymerman A., Stankiewicz A., Kolon K., and Kempers A.J. 2009. Self-organizing feature map (neural networks) as a tool to select the best indicator of road traffic pollution (soil, leaves or bark of

- Robinia pseudoacacia* L.), Environmental Pollution 157: 2061–2065.
- 32- Samecka-Cymerman A., Stankiewicz A., Kolon K., and Kempers A.J. 2007. Self-organizing feature map (neural networks) as a tool in classification of the relations between chemical composition of aquatic bryophytes and types of streambeds in the Tatra national park in Poland. Chemosphere 67: 954–960.
- 33- Vakil-Baghmisheh M.T., and Pavešić N. 2003. Premature clustering phenomenon and new training algorithms for LVQ4A2. Pattern Recognition,36:1901-1921.
- 34- Vakil-Baghmisheh M.T. 2002. Farsi Character Recognition Using Artificial Neural Networks. PhD Thesis, Faculty of Electrical Engineering, University of Ljubljana.
- 35- Vega F.A., Matias J.M., Andrade M.L., Reigosa M.J., and Covelo E.F. 2009. Classification and regression trees (CARTs) for modelling the sorption and retention of heavy metals by soil. Journal of Hazardous Materials 167: 615–624.
- 36- Yetilmezsoy K., and Demirel S. 2008. Artificial neural network (ANN) approach for modeling of Pb(II) adsorption from aqueous solution by Antep pistachio (*Pistacia Vera* L.) shells. Journal of Hazardous Materials 153: 1288–1300.



Evaluation of the Ability of LVQ4a2 Artificial Neural Network Model to Predict the Spatial Distribution Pattern of Cadmium in Soil

H. Ghorbani^{1*} · A. Roohani² · N. Hafezi Moghaddas³

Received: 14-03-2012

Accepted: 30-12-2012

Abstract

In this research, a learning vector quantization neural network (LVQ) model was developed to predict and classify the spatial distribution of cadmium in soil in Golestan province. The cadmium data were obtained from soils measuring total Cd contents in soil samples. Some statistical tests, such as means comparision, variance and statistical distribution were performed between the observed points samples data and the estimated cadmium values to evaluate the performance of the pattern recognition method. The Results showed that in training and test phase, there were no significant differences, with the confidence level of 95%, between the statsitical parameters such as average, variance, statistical distribution and also coefficient of determination in the observed and the estimated cadmium concentrations. The results suggest that learning vector quantization (LVQ) neural network can learn cadmium cocentration model precisely. In addition the results also indicated that trained LVQ neural network had a high capability in predicting cadmium concentrations for non-sampled points. The technique showed that the LVQNN could predict and map the spatial cadmium concentrations variability. Our results indicated that it is possible to discriminate different cadmium levels in soil, using LVQNN.

Keywords: Neural Network, Classification, Cadmium, Soil, Golestan Province

1, 2- Assistant Professors, Department of Water and Soil, Faculty of Agriculture, Shahrood University of Technology, Shahrood, Iran

(*-Corresponding Author Email: ghorbani1969@yahoo.com)

3- Associate Professor Faculty of Sciences, Ferdowsi University of Mashhad, Iran