بررسی کارایی مدل شبکه عصبی مصنوعی در پیش‌بینی مقدار پروتئین اندام هوایی دو رقم گندم (*Triticum aestivum* L.) تحت تنش کادمیوم

ایمان جوادزرین [۱] و بابک متشرع زاده [۲]
۱- دانش‌آموخته کارشناسی‌ارشد، گروه علوم و مهندسی خاک پردیس کشاورزی و منابع طبیعی دانشگاه تهران.۲ -دانشیار گروه علوم و مهندسی خاک پردیس کشاورزی و منابع طبیعی دانشگاه تهران.

چکیده
هدف از این پژوهش، بررسی امکان استفاده از مدل شبکه عصبی به منظور پیش‌بینی مقدار پروتئین در اندام هوایی دو رقم گندم تحت تنش کادمیوم بود. تیمارهای آزمایش شامل چهار سطح کادمیوم (صفر (تیمار شاهد)، ۲۵، ۵۰ و ۱۰۰ میلی‌گرم کادمیوم بر کیلوگرم خاک) و دو رقم گندم (آزادی و الوند) بود. از بین ۱۹ پارامتر مورد بررسی، تعداد ۱۰ پارامتر شامل مقدار کادمیوم، آهن، مس، منگنز و روی در ریشه و اندام هوایی که بیشترین همبستگی را با مقدار پروتئین اندام هوایی نشان دادند، به عنوان ورودی‌های مدل شبکه عصبی انتخاب شدند. نتایج نشان داد که مدل‌های بهینه، جهت پیش بینی مقدار پروتئین برای ارقام آزادی و الوند زمانی به دست آمد که تعداد نورون‌های لایه مخفی به ترتیب ۳ (رقم آزادی) و ۲ (رقم الوند) عدد بود. مدل شبکه عصبی بهینه شده در پیش‌بینی مقدار پروتئین اندام هوایی در ارقام آزادی (R^2=۰.۹۹) و الوند (R^2=۰.۸۴) کارایی بسیار بالایی داشت.
واژه های کلیدی: گندم، کادمیوم، شبکه عصبی، پروتئین، آلودگی خاک.

مقدمه
برای بیش از ۴۰ سال در ایران کود سوپرفسفات به مقدار ۱۰۰ تا ۳۰۰ کیلوگرم در هکتار مصرف شده است که مقدار کادمیوم استخراج شده توسط اسید نیتریک به ازای هر کیلوگرم فسفر حاصل از این کود، معادل ۶/۷ میلی‌گرم بوده است (Jalali and Khanlari, ۲۰۰۸). کادمیوم ممکن است برداشت و تجمع عناصر غذایی توسط گیاهان را با اثر گذاری بر نفوذ پذیری غشای پلاسمایی مختل کرده و منجر به کمبود یا عدم توازن عناصر غذایی در بخش های مختلف گیاهان شود (Wang-da et al., ۲۰۰۹). یکی از راهکارها برای درك توانایی گیاهان زراعی در تحمل تنش‌های غیرزیستی، شناسایی تغییرات القا شده با تنش در مقدار پروتئین این گیاهان است (Kawasaki et al. ۲۰۰۱). روش‌های اندازه‌گیری مقدار پروتئین معمولا از نظر زمان و هزینه به صرفه نیستند. یک روش جایگزین، استفاده از مدل‌های پیش‌بینی کننده از جمله شبکه عصبی مصنوعی است. یک مدل شبکه عصبی از نوع پرسپترون چند لایه (MLP) به طور معمول شامل سه لایه‌ی ورودی، مخفی و خروجی است، در هر لایه یک یا چند عنصر پردازشگر (نورون) وجود دارد که عملکرد آن‌ها شبیه به عملکرد نورون‌های مغز انسان است (Keshavarzi and Sarmadian, ۲۰۱۰). هر نورون شامل دو بخش است، در بخش اول مجموعه وزن‌دار شده، مقادیر ورودی محاسبه شده و در بخش دوم این مقادیر به عنوان ورودی در یک تابع ریاضی که تابع انتقال نامیده می‌شود، قرار گرفته و خروجی نورون محاسبه می‌شود (Keshavarzi and Sarmadian, ۲۰۱۰). شبکه عصبی مصنوعی در تحقیقات مرتبط با علوم خاک به طور گسترده‌ای استفاده شده است (Buszewski and Kowalkowski, ۲۰۰۶). توجه به تمامی مسائل مرتبط با گندم، به عنوان یک محصول راهبردی در سبد غذایی خانوارهای ایرانی با هدف تامین امنیت غذایی جامعه اهمیت دارد. هدف از این تحقیق، بررسی کارایی مدل شبکه عصبی مصنوعی در پیش بینی مقدار پروتئین اندام هوایی دو رقم گندم تحت تنش کادمیوم بود.

مواد و روش‌ها
برای انجام این تحقیق، آزمایشی فاکتوریل بر پایه طرح کاملا تصادفی در سه تکرار در گلخانه تحقیقاتی پردیس کشاورزی و منابع طبیعی دانشگاه تهران اجرا شد. تیمارهای آزمایش شامل چهار سطح کادمیوم (صفر Cd. (شاهد)، ۲۵ ($Cd_{۲۵}$)، ۵۰ ($Cd_{۵۰}$) و ۱۰۰ ($Cd_{۱۰۰}$) میلی‌گرم کادمیوم بر کیلوگرم خاک) و دو رقم گندم (آزادی و الوند) بود. تیمارهای کادمیوم از طریق اسپری کردن محلول نمک نیترات کادمیوم ($Cd(NO_۳)_۲ \cdot ۲H_۲O$) به خاک گلدان‌ها اعمال شد. تمام گلدان‌ها قبل از کشت بذر، به مدت شش ماه به منظور ایجاد تعادل در خاک در شرایط انکوباسیون نگه‌داری شدند. جهت ضد عفونی کردن بذرها، ابتدا بذرها به مدت ۳۰ ثانیه در الکل اتیلیک ۹۶ درصد غوطه ور شده سپس برای ۱۵ ثانیه در محلول آب اکسیژنه قرار گرفته و در نهایت با آب مقطر شسته شدند. قبل از کشت، خاک زراعی از مزرعه پردیس کشاورزی و منابع طبیعی دانشگاه تهران تهیه و ویژگی‌های فیزیکی و شیمیایی آن بر اساس روش‌های استاندارد تعیین گردید (امامی، ۱۳۷۵) (جدول ۱).

جدول ۱- برخی ویژگی های فیزیکی و شیمیایی خاک مورد آزمایش

خصوصیت	مقدار
کلاس بافت خاک	لومی رسی

اسیدیته خاک	۱۰/۸
قابلیت هدایت الکتریکی (dSm^{-1})	۵۵/۱
ظرفیت تبادل کاتیونی ($Cmol_c kg^{-1}$)	۲/۱۲
کربنات کلسیم معادل (درصد)	۰۰/۸
پتاسیم (میلی گرم بر کیلوگرم)	۰۰/۱۷۳
فسفر (میلی گرم بر کیلوگرم)	۰۰/۱۷
مس (میلی گرم بر کیلوگرم)	۳۸/۴
آهن (میلی گرم بر کیلوگرم)	۶۶/۵
منگنز (میلی گرم بر کیلوگرم)	۳۳/۳
روی (میلی گرم بر کیلوگرم)	۶۹/۲
کادمیوم (میلی گرم بر کیلوگرم)	۷۵/۰

واحدهای آزمایش شامل گلدان‌های سه کیلوگرمی بود. تعداد پانزده بذر در هر گلدان کشت شد که پس از ده روز به تعداد ۱۰ گیاهچه در هر گلدان کاهش یافت. شرایط دمایی گلخانه بین ۲۵ تا ۳۰ درجه سلسیوس حفظ شد. آبیاری گلدان‌ها به مدت ۳۰ روز با آب مقطر به روش وزنی صورت گرفت. نمونه‌برداری ۳۰ روز پس از کشت بذر صورت گرفت که مطابق با مقیاس زادوکس (decimal code۳۱,۳۲,۳۳,۳۴) تقریبا هم‌زمان با مرحله به ساقه رفتن گندم بود (Zadoks et al., ۱۹۷۴). مقدار عناصر کادمیوم، آهن، روی، منگنز و مس در اندام هوایی و ریشه با استفاده از دستگاه جذب اتمی (شیمادزو AA-۶۸۰۰) اندازه‌گیری شد (امامی، ۱۳۷۵). همچنین اقدام به اندازه‌گیری مقدار کلروفیل a و b (Arnon, ۱۹۴۹)، سطح برگ (leaf area-meter model LI-۳۰۰۰)، ارتفاع بوته، وزن تر و خشک اندام هوایی و ریشه و محتوی نسبی آب (Barr, H.D. and Weatherley, ۱۹۶۲) شد.

جهت انتخاب نمونه مناسب برای تعیین مقدار پروتئین، از هر واحد آزمایشی، یک بوته انتخاب شد که وضعیتی مشابه از لحاظ میزان رشد و موقعیت آن در اکثر گلدان‌ها داشت. سپس برگ دوم و سوم از بالا که رشد رویشی کافی داشت، برداشت شد و بلافاصله در نیتروژن مایع منجمد شد. نمونه‌ها بلافاصله به فریزر با دمای ۸۰- درجه سانتی‌گراد منتقل شدند. برای اندازه‌گیری مقدار پروتئین، ۵/۰ گرم از نمونه‌های منجمد شده در یک هاون ریخته شد و سپس دو میلی‌لیتر بافر فسفات (بافر استخراج) ۱۰۰ میلی‌مولار (با اسیدیته ۸/۷) که محتوی EDTA ۱/۰ میلی‌مولار و پلی وینیل پیرو لیدون (PVPP) یک درصد بود، اضافه شد و در ادامه با استفاده از نیتروژن مایع محتویات هاون ساییده شد تا یک مخلوط همگن تهیه شود. محتویات هاون به یک فالکون ۱۵ سی سی منتقل شد و توسط دستگاه سانتریفیوژ برای مدت ۲۰ دقیقه با سرعت ۱۲ هزار دور در دقیقه در دمای چهار درجه سانتی گراد سانتریفیوژ شد (Cho and Seo, ۲۰۰۵). استخراج همه نمونه‌ها در مدت کوتاهی انجام و همواره از بافر تازه استفاده شد. پس از سانتریفیوژ نمودن نمونه‌ها، مایع رویی داخل فالکون های ۱۵ سی سی در پنج تیوب اپندورف ۵/۱ سی سی تقسیم شد و بلافاصله به تانکر نیتروژن مایع منتقل گردید. برای اندازه‌گیری پروتئین نمونه‌ها در عصاره‌های استخراج شده، مقدار ۵ میکرولیتر از پروتئین استخراج شده به ۲۹۵ میکرولیتر محلول برادفورد اضافه شد و پس از به هم زدن کامل، پس از ۱۵ دقیقه میزان جذب مخلوط در طول موج ۵۹۵ نانومتر با استفاده از دستگاه الیزا ریدر (Coastar-UV Plate) قرائت شد. برای تعیین غلظت پروتئین، از سرم گاوی (BSA) به عنوان استاندارد استفاده شد (Bradford, ۱۹۷۶).

به منظور مدل‌سازی، در گام اول، ضرایب همبستگی بین ۱۹ پارامتر اندازه گیری شده (شامل غلظت عناصر کادمیوم، آهن، روی، منگنز و مس در اندام هوایی و ریشه، مقدار کلروفیل a و b، وزن تر و خشک اندام هوایی و ریشه، محتوی نسبی آب، ارتفاع بوته و سطح برگ) و مقدار پروتئین تعیین شد. سپس با توجه به ضرایب همبستگی از بین ۱۹ پارامتر مورد بررسی، ۱۰ پارامتر شامل مقدار کادمیوم، آهن، مس، منگنز و روی در ریشه و اندام هوایی که بیشترین همبستگی را با مقدار پروتئین داشتند، انتخاب شدند (جدول ۲). پارامترهای ده گانه به عنوان ورودی‌های مدل شبکه عصبی مصنوعی برای پیش‌بینی مقدار پروتئین در نظر گرفته شدند.

جدول ۲- ضرایب همبستگی بین مقدار پروتئین و پارامترهای ده گانه

۰.۷۸	پروتئین	کادمیوم اندام هوایی
۰.۷۶	پروتئین	کادمیوم ریشه
-۰.۶۶	پروتئین	مس اندام هوایی
-۰.۶۲	پروتئین	مس ریشه
-۰.۷۳	پروتئین	آهن اندام هوایی
-۰.۷۱	پروتئین	آهن ریشه
-۰.۶۴	پروتئین	منگنز اندام هوایی
-۰.۶۳	پروتئین	منگنز ریشه
-۰.۵۸	پروتئین	روی اندام هوایی
-۰.۶	پروتئین	روی ریشه

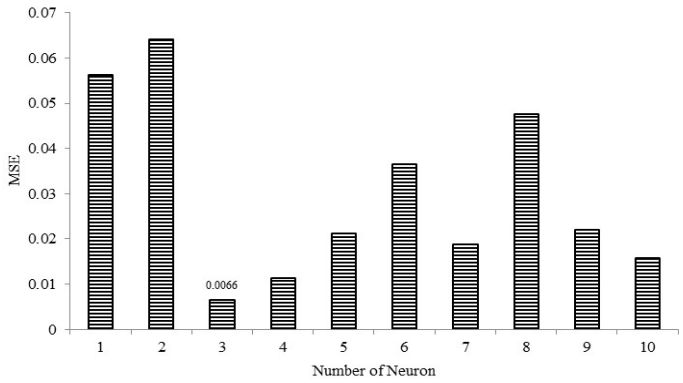
جهت ارزیابی کارایی مدل طراحی شده از شاخص میانگین مربعات خطا استفاده شد. میانگین مربعات خطا شاخصی دقیق و قابل اطمینان جهت واسنجی و آزمودن مدل شبکه عصبی است و به این صورت تعریف می‌شود (Keshavarzi and Sarmadian, ۲۰۱۰):

$$MSE = \frac{1}{n}\sum_{k=1}^{n}(Z_o - Z_p) \quad (1)$$

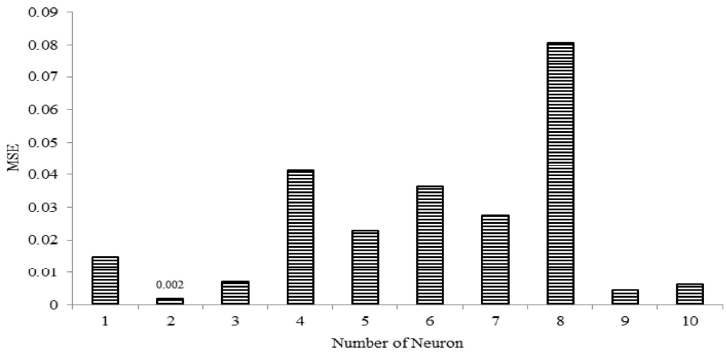
در این معادله Z_o مقدار اندازه‌گیری شده، Z_p مقدار پیش‌بینی شده و n تعداد نمونه‌های آموزشی است. در این تحقیق از الگوریتم لونبرگ-مارکوئت (LM) به عنوان الگوریتم آموزشی استفاده شد. توابع انتقالی لایه‌های مخفی و خروجی به ترتیب تانژانت سیگموئید و پیورلاین انتخاب شدند. جهت تعیین بهترین معماری شبکه عصبی، نورون‌های شبکه به تدریج تا ۱۰ نورون در لایه مخفی افزایش یافت و به ازای افزودن هر نورون، شبکه طراحی شده تا ۵ مرتبه اجرا شد و در ادامه اقدام به محاسبه میانگین MSE حاصل از ۵ تکرار هر یک از مدل‌ها شد. کمترین مقدار میانگین MSE به عنوان بهترین معماری شبکه عصبی مصنوعی انتخاب شد. به طور کلی برای هریک از ارقام آزادی و الوند یک مدل جداگانه طراحی شد. پس از آموزش شبکه و به منظور بررسی کارایی مدل انتخاب شده، اقدام به شبیه‌سازی مدل شبکه عصبی شد، به این صورت که از مجموعه‌ای از داده‌های آزمایشی که در مرحله آموزش شبکه به کار برده نشده بود، به عنوان مجموعه داده‌های ورودی برای بهترین مدل شبکه عصبی مصنوعی استفاده شد. در مرحله شبیه سازی، یک مدل شبکه عصبی، داده‌های خروجی را با توجه به مقادیر داده‌های ورودی ایجاد می‌کند. در نهایت نتایج حاصل از مدل شبکه عصبی مصنوعی با نتایج واقعی مقایسه شد.

نتایج و بحث

مدل بهینه شبکه عصبی مصنوعی جهت پیش‌بینی مقدار پروتئین برای ارقام آزادی و الوند زمانی به دست آمد که تعداد نورون‌های لایه مخفی به ترتیب ۳ و ۲ عدد بود (شکل‌های ۱ و ۲).

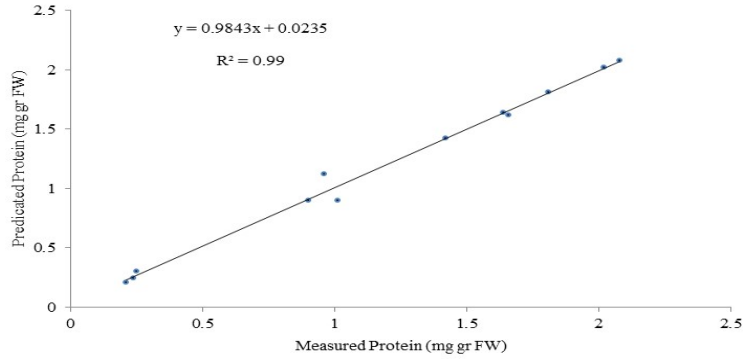


شکل ۱- تغییرات مقدار MSE با تغییر تعداد نورون های لایه مخفی مدل های شبکه عصبی مصنوعی رقم آزادی

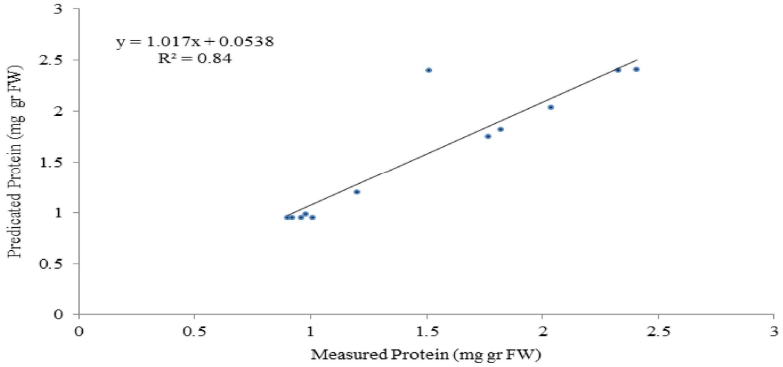


شکل ۲- تغییر مقدار MSE با تغییر تعداد نورون های لایه مخفی مدل های شبکه عصبی مصنوعی رقم الوند

در رابطه با رقم آزادی، زمانی که تعداد نورون‌های لایه مخفی به ۳ نورون افزایش یافت، مقدار MSE کاهش قابل توجهی نشان داد. به منظور اطمینان از این مسئله که مقدار ملاحظه شده MSE یک مینیموم محلی نیست، تعداد نورون‌های لایه مخفی تا ۱۰ نورون افزایش یافت. هر چند زمانی که تعداد نورون‌های لایه مخفی ۴ و ۱۰ نورون بود، مقدار MSE کاهش یافته بود اما در مقایسه با زمانی که ۳ نورون در لایه مخفی وجود داشت، تعداد نورون‌های ۴ و ۱۰ در لایه مخفی نشان دهنده مینیم‌های محلی بودند، بنابراین مقدار MSE مدل شبکه عصبی مصنوعی با ۳ نورون در لایه مخفی نشان دهنده یک مینیمم مطلق بود. بنابراین مدل شبکه عصبی با ۳ نورون در لایه مخفی به عنوان مدل بهینه برای پیش‌بینی مقدار پروتئین اندام هوایی در رقم آزادی معرفی شد (شکل ۱). بر اساس استدلالی مشابه، مدل بهینه برای پیش بینی مقدار پروتئین اندام هوایی رقم الوند زمانی به دست آمد که ۲ نورون در لایه مخفی شبکه عصبی مصنوعی وجود داشت (شکل ۲). شکل های ۳ و ۴ به ترتیب نتایج حاصل از مدل‌های بهینه شبکه عصبی مصنوعی را در پیش‌بینی مقدار پروتئین برای ارقام آزادی و الوند نشان می‌دهد.



شکل ۳- نمودار پراکندگی مقدار اندازه گیری شده و پیش بینی شده پروتئین اندام هوایی در رقم ازادی



شکل ۴ - نمودار پراکندگی مقدار اندازه گیری شده و پیش بینی شده پروتئین اندام هوایی در رقم الوند

نتایج حاصل از شبیه‌سازی مدل بهینه مربوط به رقم آزادی در شکل ۳ نشان داده شده است. این شکل نشان می‌دهد زمانی که از پارامترهای ده‌گانه منتخب به عنوان بردارهای ورودی برای مدل شبکه عصبی مصنوعی بهینه در مرحله آموزش مدل استفاده شود، این مدل کارایی بسیار بالایی (R^2=۰.۹۹) در مرحله شبیه‌سازی جهت پیش‌بینی مقدار پروتئین اندام هوایی رقم آزادی داشت (شکل ۳). هر چند مقدار MSE مدل بهینه رقم الوند در مرحله آموزش کمتر از مقدار MSE رقم آزادی بود (به عبارت دیگر در مرحله آموزش شبکه عصبی مصنوعی، دقت مدل بهینه الوند بیش از مدل بهینه آزادی بود)، اما در مرحله شبیه‌سازی مدل‌ها، کارایی مدل بهینه رقم آزادی به طرز قابل ملاحظه‌ای بیش از رقم الوند بود. نورون‌ها واحدهای اصلی پردازش مدل‌های شبکه عصبی مصنوعی هستند، بنابراین به نظر می‌رسد تعداد نورون‌های لایه مخفی یک مدل بهینه شبکه عصبی در مرحله آموزش شبکه در مقایسه با مقدار MSE آن مدل، عاملی تعیین‌کننده در میزان کارایی مدل بهینه شبکه عصبی مصنوعی در مرحله شبیه‌سازی است. کارایی مدل بهینه مربوط

به رقم الوند در مقایسه با رقم آزادی به منظور پیش‌بینی مقدار پروتئین کمتر بود (R^2=۰.۸۴). شیب نمودار پراکندگی و دامنه تغییرات پروتئین در رقم الوند بیش از رقم آزادی بود (شکل ۴ در مقایسه با شکل ۳)، به عبارت دیگر در رقم الوند علاوه بر این که تغییرات مقدار پروتئین شدیدتر اتفاق افتاده است، این تغییرات در دامنه وسیع‌تری نیز در مقایسه با رقم آزادی اتفاق افتاده است. می‌توان از این مسئله استدلال کرد که کارایی سیستم آنتی‌اکسیدان در رقم الوند بیش از رقم آزادی است، هر چند اثبات این موضوع نیاز به بررسی آن دارد.

منابع

امامی، ع؛ (۱۳۷۵). روش های تجزیه گیاه.موسسه تحقیقات خاک و آب، نشریه فنی شماره ۹۸۲، جلد اول.

Arnon DI. ۱۹۴۹. copper enzymes in isolated chloroplasts, polyphenoxidase in beta vulgaris. Plant physiology ۲۴: ۱-۱۵.

Barr, H.D. and Weatherley, P.E. (۱۹۶۲). A re-examination of the relative turgidity technique for estimating water deficit in leaves. Aust. J. Biol. Sci. ۱۵:۴۱۳-۴۲۸.

Bradford, M.M., ۱۹۷۶. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of proteindye binding. Anal. Biochem. ۷۲, ۲۴۸-۵۴.

Cho, U.H., Seo, N.H., ۲۰۰۵ Oxidative stress in Arabidopsis thaliana exposed to cadmium is due to hydrogen peroxide accumulation. Plant Sci. ۱۶۸, ۱۱۳-۱۲۰.

Jalali, M and Khanlari, Z. V. (۲۰۰۸). Cadmium Availability in Calcareous Soils of Agricultural Lands in Hamadan, Western Iran. Soil and Sediment Contamination, ۱۷: ۲۵۶-۲۶۸.

Wang-da, C. Hai-gen, Y. Hong-mei, Z. Xian-guo, T. (۲۰۰۹). Influences of Cadmium on Grain Mineral Nutrient Contents of Two Rice Genotypes Differing in Grain Cadmium Accumulation. Rice Science, ۱۶(۲): ۱۵۱-۱۵۶.

Kawasaki, S. and Borchert, C. ۲۰۰۱. Gene expression profiles during the initial phase of salt stress in Rice. The Plant Cell ۱۳: ۸۸۹-۹۰۵.

Keshavarzi A. and F. Sarmadian. (۲۰۱۰). Comparison of Artificial Neural Network and Multivariate Regression Methods in Prediction of Soil Cation Exchange Capacity. International Journal of Environmental and Earth Sciences ۱:۱.

Buszewski, B. Kowalkowski, T. (۲۰۰۶). A new model of heavy metal transport in the soil using non-linear artificial neural networks. Environ. Eng. Sci. ۲۳ (۴): ۵۸۹-۵۹۵.

Zadoks, J.C. Chang, T.T. Konzak, C.F. "A Decimal Code for the Growth Stages of Cereals", Weed Research ۱۹۷۴ ۱۴:۴۱۵-۴۲۱.

Abstract

The aim of this experiment was to evaluate the possibility of using neural network model to predict content protein in shoot of two cultivars wheat under cadmium stress. The treatments consisted of four levels of Cd (۰ (as control, Cd.), ۲۵ ($Cd_{۲۵}$), ۵۰ ($Cd_{۵۰}$) and ۱۰۰ ($Cd_{۱۰۰}$) mg Cd/kg soil) and two wheat cultivars (Azadi and Alvand), respectively. ۱۰ parameters of ۱۹ investigated parameters (including amount of cadmium, iron, copper, manganese and zinc in roots and shoots that showed the highest correlation with amount of protein of shoot) selected as inputs to the neural network model. Results showed that optimum models for predicting amount of protein in shoot of studied cultivars was obtained when the numbers of hidden layer neurons were ۳ (Azadi) and ۲ (Alvand) respectively. Optimized neural network model had very high efficiency to predicate the amount of protein in shoot of Azadi ($R^۲$ = ۰.۹۹) and Alvand ($R^۲$ = ۰.۸۴) respectively.