پانزدهمین کنگره علوم خاک ایران
۶ تا ۸ شهریور ۱۳۹۶ محور مقاله: ارزیابی و تناسب اراضی

مقایسه روش‌های شبکه عصبی مصنوعی و رگرسیونی در برآورد عملکرد چای بر اساس ویژگی‌های خاکی

نفیسه یغمائیان مهابادی[1*]، کسری سمیعی[2]
به ترتیب استادیار و دانش‌آموخته کارشناسی‌ارشد گروه علوم خاک، دانشکده کشاورزی، دانشگاه گیلان

چکیده

هدف از این مطالعه مقایسه کارآیی روش‌های شبکه عصبی مصنوعی و رگرسیون چند متغیره به منظور تخمین عملکرد چای بر اساس ویژگی‌های خاکی می‌باشد. به این منظور در باغی به وسعت حدود ۳۰ هکتار در منطقه فومن استان گیلان، ۷۰ نمونه خاک از عمق صفر تا ۵۰ سانتی‌متر برداشت شد. به منظور تعیین متوسط عملکرد چای، نمونه‌برداری از برگ سبز چای به مرکزیت محل‌های نمونه‌برداری خاک انجام شد. در نهایت مدل‌های رگرسیونی و شبکه عصبی مصنوعی بین ویژگی‌های خاک و عملکرد برقرار و سپس اعتبارسنجی شدند. نتایج مطالعه نشان داد که مدل شبکه عصبی مصنوعی توانست حدود ۷۹ درصد از تغییرات مکانی عملکرد و مدل رگرسیونی حدود ۵۹ درصد از تغییرات عملکرد در منطقه مورد مطالعه را توجیه نماید. نتایج آنالیز حساسیت مدل شبکه عصبی نشان داد که درصد رس، کربن آلی و pH خاک به عنوان مهم‌ترین پارامترهای مؤثر بر پیش‌بینی تولید چای در منطقه مورد مطالعه می‌باشند. مدل شبکه عصبی با ME و RMSE به ترتیب ۰/۰۱۶ و ۰/۱۸۱ از کارآیی بالاتری در تخمین عملکرد چای نسبت به مدل رگرسیونی برخوردار است.

واژه های کلیدی: تخمین عملکرد، آنالیز حساسیت، رگرسیون خطی چند متغیره

مقدمه

برای پیش‌بینی مؤثر عملکرد گیاهان زراعی، ابتدا باید عوامل تعیین‌کننده یا محدودکننده عملکرد را شناخت. تولیدات کشاورزی همیشه با محدودیت شرایط اقلیمی، خاکی و تغییرات بازارهای بین‌المللی همراه بوده است. اگرچه این احتمال خطر، هرگز به طور کامل حذف نمی‌شود، اما می‌توان با شناخت عوامل مختلف مؤثّر در رشد گیاه و محصول و تخمین میزان محصول قبل از فصل برداشت، آنها را به حداقل رساند (Qian et al., 2009). عملکرد گیاه تابعی از عوامل مختلف گیاهی، اقلیمی و شرایط مدیریتی آب و خاک است. از اینرو محاسبه مقدار عملکرد گیاه و شاخص‌های وابسته به آن از روابط غیر خطی پیچیده‌ای پیروی می‌کند که مدلسازی آن نیز دشواری خاصی دارد. با توجه به اهمیت مدل‌سازی عملکرد و تعیین عوامل مؤثر بر رشد، امروزه مدل‌های شبیه‌سازی گیاهی به عنوان ابزاری چندمنظوره در تحقیقات مطرح هستند. یکی از روش‌های مدل‌سازی که در سال‌های اخیر مورد توجه بسیاری از محققین در علوم مختلف واقع شده است، مدل‌سازی به روش شبکه عصبی مصنوعی می‌باشد. علاوه بر شبکه عصبی مصنوعی، مدل‌های رگرسیونی نیز از روش‌های رایج در پیش‌بینی عملکرد محصول می‌باشند. شبکه عصبی مصنوعی جزئی از سیستم‌های دینامیکی هوشمند هستند که با پردازش داده‌های مشاهداتی، قانون نهفته در ورای اطلاعات را به ساختار شبکه منتقل می‌کنند. به تعبیر دیگر شبکه عصبی مصنوعی، مدلی ریاضی است که توانایی مدل‌سازی و ایجاد روابط غیر خطی برای درون‌یابی را دارد. در این روش تلاش بر این است که بر اساس روابط ذاتی میان داده‌ها، نگاشتی غیر خطی بین متغیرهای مستقل و وابسته برقرار گردد (Kaul et al., 2005).

دراموند و همکاران (Drummond et al., 2003) در پژوهش خود جهت تعیین روابط میان خصوصیات خاک، توپوگرافی و عملکرد غلات در منطقه میسوری آمریکا مشاهده نمودند که روش شبکه‌های عصبی مصنوعی نسبت به روش‌های رگرسیونی دارای خطای کم‌تری می‌باشد. در مطالعه‌ای دیگر از قابلیت شبکه‌های عصبی مصنوعی برای پیش‌بینی عملکرد ذرت و سویا در منطقه مریلند استفاده شده و نتایج حاکی از آن بود که مدل شبکه عصبی مصنوعی در مقایسه با مدل‌های رگرسیونی دارای

دقت بالاتری می‌باشد (Kaul et al., 2005). آلوارز (Alvarez, 2009) شبکه‌های عصبی مصنوعی را جهت تعیین عملکرد گندم در منطقه پامپاس آرژانتین به کار برد. در این تحقیق نسبت بارندگی به تبخیر و تعرق (R/CPET) به عنوان مهم‌ترین فاکتور آب و هوایی مؤثر بر بازده این محصول معرفی شد. صفا و همکاران (Safa et al., 2015) بیان کردند که شبکه‌های عصبی مصنوعی با ورودی‌های شرایط مزرعه و نهاده‌های کشاورزی شامل سطح زیر کشت، دور آبیاری، میزان مصرف کود ازت و قارچ‌کش‌ها قادر به تخمین عملکرد گندم در منطقه کانتربوری نیوزلند می‌باشد. اسمعیل‌زاده (۱۳۹۱)، در بررسی مقادیر به دست آمده از شاخص های آماری میانگین خطا، میانگین خطای مطلق، جذر میانگین مربعات خطا، کفایت مدل و جذر میانگین مربعات خطای استاندارد برای دو مدل شبکه عصبی مصنوعی و رگرسیون چندگانه نشان داد که مدل نوین شبکه عصبی مصنوعی کارآیی بیش‌تری در تخمین مکانی عملکرد برنج نسبت به روش رگرسیون چندگانه دارد.

با توجه به اینکه کشت و صنعت چای نقش بسیار مهمی را در گردش اقتصادی مناطق چایکاری شمال کشور بر عهده دارد؛ برآورد عملکرد این محصول نقش مهمی در سیاست‌گذاری‌های بخش کشاورزی ایفا می‌کند. همچنین شناخت ویژگی‌های خاکی تأثیرگذار بر روی عملکرد چای جهت برنامه‌ریزی و اعمال مدیریت بهینه حائز اهمیت می‌باشد. این پژوهش با هدف بررسی کارآیی روش‌های شبکه عصبی مصنوعی و رگرسیون چند متغیره به منظور تخمین عملکرد چای بر اساس ویژگی‌های خاک در منطقه فومن استان گیلان انجام گرفت.

مواد و روش‌ها

منطقه مورد مطالعه یک باغ چای به وسعت حدود ۳۰ هکتار واقع در ۱۰ کیلومتری غرب شهرستان فومن استان گیلان، بین طول‌های جغرافیایی ″۰۲ ′۱۳ °۴۹ و ″۲۴ ′۱۳ °۴۹۰ و عرض‌های جغرافیایی ″۳۱ ′۱۳ °۳۷۰ و ″۰۴ ′۱۴ °۳۷۰ قرار دارد. این منطقه دارای میانگین دمای سالیانه ۱۵/۷ درجه سانتی‌گراد و بارندگی سالیانه ۱۲۷۵ میلی‌متر می‌باشد. متوسط ارتفاع منطقه از دریا، ۳- متر و سیمای اراضی دشت رودخانه‌ای می‌باشد. بر اساس الگوی نمونه‌برداری شبکه‌ای منظم با ابعاد ۷۵ در ۷۵ متر، مختصات جغرافیایی ۷۰ نقطه مشاهداتی به دست آمد. نمونه‌برداری خاک از عمق صفر تا ۵۰ سانتی‌متر از نقاط با مختصات معلوم و در فاصله بین ردیف‌های چای انجام گرفت. نمونه‌های خاک، هوا خشک شده و بعد از کوبیده شدن از الک ۲ میلی‌تر عبور داده شدند. سپس pH در عصاره ۱:۲/۵ خاک به آب توسط دستگاه pH متر، هدایت الکتریکی عصاره اشباع توسط دستگاه هدایت‌سنج، پتاسیم قابل استفاده به روش عصاره‌گیری با استات آمونیوم، فسفر قابل استفاده به روش بیکربنات سدیم، نیتروژن کل خاک به روش کجلدال، درصد کربن آلی به روش والکی بلاک، درصد اجزای معدنی خاک به روش هیدرومتری، جرم مخصوص ظاهری به روش کلوخه و پارافین و درصد رطوبت اشباع در نمونه‌های خاک اندازه‌گیری شد (Soil Survey Staff. 1996). برای تعیین عملکرد، نمونه‌برداری از برگ سبز چای در پلاتی به وسعت ۲ متر مربع به مرکزیت محل‌های نمونه‌برداری خاک انجام شد. به منظور تعیین متوسط عملکرد چای (وزن تر)، یک غنچه و دو برگ انتهایی چای در سه برداشت متوالی (برداشت اول در ماه اردیبهشت برداشت دوم در ماه خرداد و برداشت سوم در اواخر ماه تیر) از سطح مذکور برداشت شد. لازم به ذکر است تمام بوته‌های چای هم‌سن بوده و تحت مدیریت مشابهی قرار داشتند.

بسط و توسعه یک مدل شبکه عصبی مصنوعی مستلزم طراحی اجزای شبکه است. جهت دستیابی به اهداف موردنظر از شبکه‌های چند لایه پرسپترون (MLP) که از متداولترین شبکه در پیش‌بینی و حل مسائل غیرخطی هستند، استفاده شد. آموزش مدل‌های شبکه عصبی بر مبنای آزمون و خطا پایه‌ریزی شد. به‌طوری‌که آرایش بهینه شبکه با تغییر تعداد لایه‌های پنهان و نرون آن، نوع تابع آموزشی و تعداد تکرار مرحله آموزش به منظور محاسبه عملکرد به‌دست آمد. به این ترتیب جهت انتخاب مدل مناسب و بهینه از آماره‌های ریشه میانگین مربعات خطا (RMSE) و ضریب تبیین (R^2) استفاده گردید (Liu et al., 2001). شبکه مزبور در محیط نرم‌افزار Neurosolutions اجرا شد. به منظور تعیین مؤثرترین ویژگی‌های خاکی در تخمین عملکرد چای، آنالیز حساسیت توسط نرم افزار Neurosolutions انجام گرفت. در مرحله بعد بین ویژگی‌های خاکی و عملکرد چای آنالیز رگرسیون چند متغیره خطی به روش رگرسیون گام به گام و توسط نرم‌افزار SPSS صورت گرفت. برای تعیین صحت

و اعتبار مدل‌ها از آماره‌های میانگین خطا (ME)، RMSE و R^2 بین مقادیر اندازه‌گیری‌شده و پیش‌بینی‌شده عملکرد توسط مدل‌های شبکه عصبی و رگرسیونی استفاده شد (Norouzi et al., 2010).

نتایج و بحث

بررسی رابطه رگرسیونی بین ویژگی‌های خاکی مورد مطالعه و عملکرد چای

خلاصه آماری ویژگی‌های خاکی مطالعه شده و اطلاعات آماری عملکرد چای در منطقه مورد مطالعه در جدول ۱ نشان داده شده است. نتایج این جدول نشان می‌دهد که عملکرد چای تغییرپذیری نسبتا زیادی در منطقه دارد. بر اساس نتایج آزمون کولموگروف-اسمیرونوف، عملکرد چای از توزیع نرمال برخوردار است. به علاوه، مقادیر چولگی نیز نتایج آزمون نرمال بودن عملکرد چای را تایید می‌کند که مقادیر آن بین ۱+ تا ۱- قرار دارد. همچنین نزدیک بودن مقادیر میانگین با مقادیر میانه دلیل دیگری بر این مدعاست. مدل رگرسیونی توانسته است ۵۹ درصد از تغییرات عملکرد چای در سطح منطقه را توجیه نماید (شکل ۱-الف)؛ به عبارتی حدود ۴۰ درصد تغییرپذیری عملکرد در منطقه مورد مطالعه توسط مدل قابل تبیین نبود که دلیل این موضوع می‌تواند وجود روابط غیرخطی بین پدیده‌ها باشد که در مدل رگرسیونی در نظر گرفته نمی‌شود. مدل رگرسیونی چند متغیره بین عملکرد چای و ویژگی‌های خاک در زیر ارائه شده است. همانطور که در مدل مشاهده می‌شود رابطه بین عملکرد با مقدار رس و کربن آلی، مستقیم و با جرم مخصوص ظاهری خاک، غیرمستقیم می‌باشد.

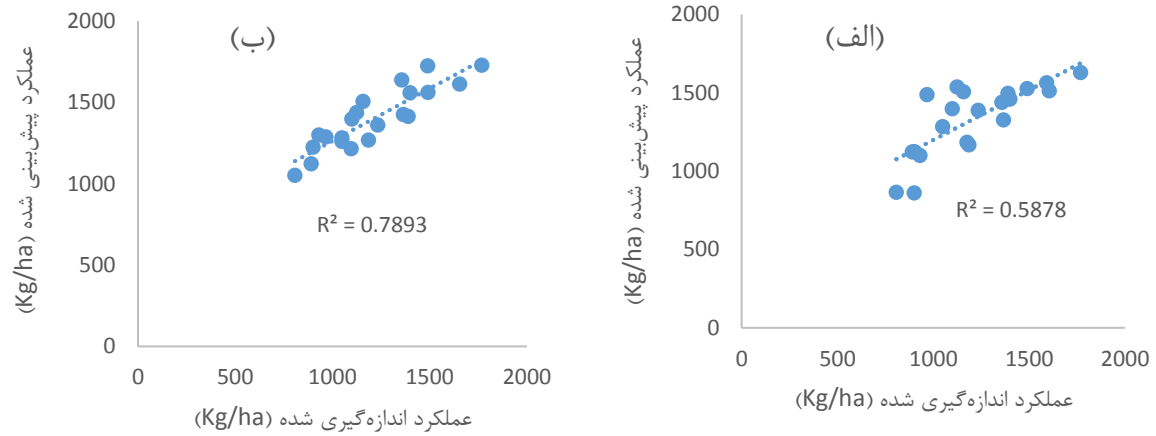
$$\text{Yield} = 3862.5 + 2100.5\ \text{clay} + 239.1\ \text{OC} - 2.86\ \rho b \qquad (۱)$$

جدول ۱- خلاصه آماری ویژگی‌های خاک و عملکرد چای

متغیر	pH	OC (٪)	EC (dS/m)	N (٪)	P (mg/kg)	K (mg/kg)	ρb (gr/cm^3)	SP (٪)	Sand (٪)	Silt (٪)	Clay (٪)	Yield (Kg/ha)
میانگین	۵/۸۴	۲/۶۳	۰/۱۶	۰/۳۲	۱۱/۸۳	۴۰۱/۴	۱/۴	۶۵/۸	۳۹/۹	۳۵	۲۴/۹	۱۴۴۲/۸
کمینه	۵/۳۵	۰/۵۹	۰/۰۸	۰/۲	۳/۶۹	۱۳۹/۹	۱/۱۳	۲۱	۱۶	۱۹/۵	۱۴/۵	۶۰۳/۸۷
بیشینه	۶/۵۹	۳/۶۷	۰/۳۷	۰/۵۷	۱۹/۷	۷۹۴/۶	۱/۶۶	۹۱	۶۶	۵۰/۵	۳۸	۲۷۷۱/۵
انحراف معیار	۰/۳	۰/۶۹	۰/۰۵	۰/۰۹	۳/۲۸	۱۴۹/۷	۰/۱۱	۱۳/۱	۷/۹۱	۵/۵۱	۴/۳۱	۴۳۷/۵
ضریب تغییرات (٪)	۵/۱	۲۶/۳	۳۱/۲	۲۸/۱	۲۷/۷	۳۷/۲	۷/۸	۱۹/۹	۱۹/۸	۱۵/۷	۱۷/۳	۳۰/۳
چولگی	۰/۴۶	-۰/۷۸	۰/۸۵	۰/۸	-۰/۲۱	۰/۵۶	۰/۰۶	-۰/۵	-۰/۱۲	۰/۵۷	۰/۵۸	۰/۵۲
کشیدگی	-۰/۴	۰/۳۹	۰/۶۵	-۰/۹۷	۰/۱۴	-۰/۲۷	-۰/۱۷	۰/۷۴	۲/۰۳	۱/۱۴	۱/۴۷	-۰/۱۴

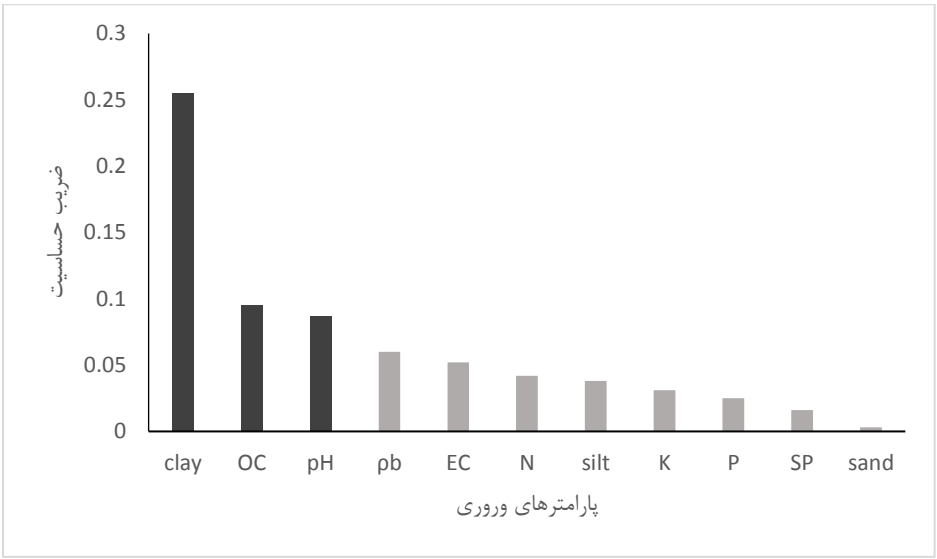
نتایج مدل‌سازی شبکه عصبی

برای آموزش شبکه عصبی مصنوعی MLP از الگوریتم آموزشی لورنبرگ-مارکاردت استفاده شد که در نهایت از روش آزمون و خطا برای تعیین بهترین معماری شبکه استفاده گردید. معماری بهینه انتخاب شده برای شبکه عصبی دارای ۱۱ نرون در لایه ورودی، ۸ نرون در لایه میانی با تابع هدف Tanhaxon و یک نرون در لایه خروجی با تابع هدف خطی با تعداد تکرار ۱۰۰۰ بار بود. مدل بهینه شبکه عصبی برای پیش‌بینی عملکرد چای در منطقه مورد مطالعه توانست ۷۹ درصد از تغییرپذیری عملکرد چای را توجیه نماید (شکل ۱-ب).



شکل ۱- رابطه بین عملکرد پیش‌بینی شده و اندازه‌گیری شده در الف: مدل رگرسیون چند متغیره و ب: مدل شبکه عصبی مصنوعی

برای شناخت ویژگی‌های خاکی مهم و مؤثر بر عملکرد چای و در نتیجه استفاده از این ویژگی‌ها به منظور دستیابی به برآورد بهتر، آنالیز حساسیت انجام شد که نتایج آن در شکل ۲ نشان داده شده است. مشاهده می‌گردد که درصد رس به عنوان مهمترین ویژگی مؤثر بر عملکرد چای در منطقه می‌باشد. به عبارتی عملکرد چای در منطقه، بیش‌ترین مقدار حساسیت را به مقدار رس خاک نشان داده است. بعد از رس بیشترین حساسیت برای کربن آلی و pH به دست آمده است. بر این اساس، درصد رس، کربن آلی و pH خاک به عنوان مهم‌ترین پارامترهای مؤثر بر پیش‌بینی تولید چای در منطقه مورد مطالعه شناخته شدند و سایر پارامترها سهم کمی در توضیح تغییرپذیری عملکرد چای در این منطقه دارند.



شکل ۲- نتایج آنالیز حساسیت بارامترهای مورد بررسی بر اساس شبکه عصبی مصنوعی

درصد ذرات معدنی خاک (رس، سیلت و شن) به‌طور مستقیم یا غیرمستقیم در تامین و نگهداری آب، هوا و عناصر غذایی مورد نیاز گیاهان نقش دارند. همچنین در خاک‌های سبک بافت و شنی، درصد بالایی از نیتروژن کودی، آبشویی شده و از دسترس ریشه گیاه خارج می‌شود (Reynolds et al., 2009). بنابراین درصد اجزای معدنی خاک به عنوان عامل تأثیرگذار بر وضعیت حاصلخیزی خاک، در میزان تولید محصول اهمیت ویژه‌ای دارد. فو و همکاران (Fu et al., 2003) گزارش کردند که ماده‌ی آلی خاک بیشترین اثر را بر توزیع جوامع گیاهی دارد. وجود ماده آلی علاوه بر این‌که نشان دهنده سلامت و کیفیت خاک است، شاخص مناسبی برای باروری خاک به شمار می‌رود که حاصل برهم‌کنش فرآیندهای فیزیکی، شیمیایی و

زیستی است و آینـده ثبـات در تولیـد محصـول بـه شـدت وابسته به بهبود شرایط حاصلخیزی خاک خواهد بود (Singh et al., 2004). از آنجا که pH بر جذب بیشتر عناصر غذایی خاک تأثیرگذار است، احتمالا به همین دلیل این ویژگی تأثیر زیادی بر عملکرد چای داشته است. افزایش pH از سویی بر فراهمی عناصر غذایی به ویژه عناصر کم‌مصرف تأثیر می‌گذارد و از طرف دیگر تغییرات pH روی جمعیت موجودات زنده خاک اثرگذار است (Christensen et al., 1995). اسمعیل‌زاده (۱۳۹۱) در بررسی میزان حساسیت عملکرد محصول برنج (به عنوان پارامتر خروجی) نسبت به ویژگی‌های خاکی (به عنوان پارامترهای ورودی مدل) به این نتیجه دست یافت که ویژگی‌های قابلیت هدایت الکتریکی عصاره‌ی اشباع خاک و اسیدیته خاک، بیش‌ترین حساسیت را بر عملکرد محصول ایجاد کرده‌اند. این در حالی است که ویژگی‌های مربوط به بافت خاک (درصد ذرات رس و شن)، تأثیر کمتری بر عملکرد محصول ایجاد کرده‌اند.

دو آماره RMSE و ME دقت مدل‌سازی در قالب تفاضل مقادیر پیش‌بینی و اندازه‌گیری شده و آماره R^2 همبستگی بین مقادیر اندازه‌گیری و پیش‌بینی شده را نشان می‌دهد. بدیهی است که نزدیک بودن مقدار آماره ME به صفر موید این مطلب است که برازش توسط مدل مورد نظر نااریب است و مقدار RMSE پایین مدل نشان‌دهنده دقت مناسب و قابل قبول برآورد می‌باشد. نزدیکی آماره R^2 به یک نیز، بیانگر همبستگی قوی بین مقادیر اندازه‌گیری و پیش‌بینی شده است (زارع ابیانه و همکاران، ۱۳۸۸). بنابراین با توجه به ME و RMSE که برای مدل شبکه عصبی به ترتیب ۰/۰۱۶ و ۰/۱۸۱ و برای مدل رگرسیونی برابر ۱۰/۸۰ و ۳۸/۶۲۳ می‌باشد، شبکه عصبی نسبت به روش رگرسیونی در پیش‌بینی عملکرد چای بر اساس ویژگی‌های خاک در منطقه مورد مطالعه دارای دقت بیشتری می‌باشد. شبکه عصبی مصنوعی یکی از موفق‌ترین تکنیک‌های داده کاوی است که از طریق یک سری مدل‌های پیوندی الهام گرفته شده از رفتار مغز انسان که دارای قابلیت‌های سودمند برای تحقیقات است، می‌تواند روابط غیرخطی بین پدیده‌ها را پیش‌بینی کند (Haykin, 1999). اگر چه روابط بین پارامترها می‌تواند به وسیله مدل‌های متنوع رگرسیونی مدل‌سازی شود ولی طبیعت و ذات این روابط طوری است که چنین مدل‌هایی نمی‌توانند چنین روابطی را به خوبی تجزیه و تحلیل کنند. شبکه عصبی مصنوعی می‌تواند روابط غیرخطی پیچیده یا لایه‌های پنهان بین متغیرهای وابسته و مستقل را بیابد و با دقت بهتری نسبت به روش رگرسیون عمل نماید (Cross et al., 1995). مطالعات دراموند و همکاران (Drummond et al., 2003) و کول و همکاران (Kaul et al., 2005) بیانگر کارآیی بیشتر مدل شبکه عصبی مصنوعی در مقایسه با مدل‌های رگرسیونی در تخمین عملکرد محصول می‌باشد.

مدل شبکه عصبی مصنوعی به دلیل در نظر گرفتن روابط غیرخطی بین عمکلرد چای و ویژگی‌های خاکی و در نتیجه افزایش دقت در تخمین‌ها می‌تواند جایگزین مدل‌های مرسوم رگرسیونی در مدل‌سازی عملکرد محصول باشد. با این وجود شبکه عصبی مصنوعی نیز با مشکلاتی از قبیل آزمون اجزای مختلف ساختار شبکه همچون تعداد لایه‌های پنهان و نرون آن، نوع توابع آموزشی، تعداد تکرار مرحله آموزش و قانون یادگیری مواجه می‌باشد.

منابع

اسمعیل‌زاده کردخیلی، س. ۱۳۹۱. برآورد عملکرد برنج با استفاده از روش‌های زمین آماری، شبکه‌های عصبی مصنوعی و رگرسیون چندگانه در شالیزارهای لشت‌نشای استان گیلان. پایان‌نامه کارشناسی ارشد علوم خاک، دانشکده‌ی علوم کشاورزی دانشگاه ولی‌عصر (عج) رفسنجان.

زارع ابیانه، ح. ۱۳۹۱. ارزیابی روش‌های شبکه عصبی مصنوعی و زمین آمار در برآورد توزیع مکانی عملکرد گندم دیم و آبی (مطالعه موردی خراسان رضوی). پژوهش‌های جغرافیایی طبیعی، شماره ۴، صفحه‌های ۲۳ تا۴۲.

Alvarez A. 2009. Predicting average regional yield and production of wheat in the Argentine Pampas by an artificial neural network approach. European Journal of Agronomy, 30:70-77.

Christensen J.V., Legge W.G., Depauw R.M., Hennig A.M.F., McKenzie J.S., Siemens B. and Thomas J.B. 1995. Effect of seeding date, nitrogen and phosphate fertilizer on growth, yield and quality of rapeseed inNorthwest Alberta. Canadian Journal of Plant Science, 65: 275-284.

Cross S.S., Harrison R.F. and Kennedy R.L. 1995. Introduction to neural networks. Lancet, 346: 1075-9.

Drummond S.T., Sudduth K.A., Joshi A., Birrell S.J. and Kitchen N.R. 2003. Statistical and neural methods for site-specific yield prediction. Transactions of the ASAE, 46: 5-14.

Fu B.J., Liu S.L., Ma K.M. and Zhu Y.G. 2003. Relationships between soil characteristic, topography and plant diversity in a heterogeneous broad-leaved forest near Beijing China. Journal of Plant and Soil, 261: 47-54.
Haykin S. 1999. Neural Networks: A Comprehensive Foundation. Prentice-Hall Inc., NJ.
Kaul M., Hill R.L. and Walthall C. 2005. Artificial neural networks for corn and soybean yield prediction. Agricultural Systems, 85: 1-18.
Liu J., Goering C.E. and Tian L. 2001. A neural network for setting target yields. Transactions of the American Society of Agricultural Engineers, 44: 705-713.
Norouzi M., Ayoubi Sh., Jalalian A., Khademi H. and Dehghani A.A. 2010. Predicting rainfed wheat quality and quantity by artificial neural network using terrain and soil characteristics. Acta Agriculture Acandinavica, Section B – Soil and Plant Science, 60: 341-352.
Qian B., Jong R.D., Warren R., Chipanshi A. and Hill H. 2009. Statistical spring wheat yield forecasting for the Canadian prairie provinces. Agricultural and Forest Meteorology, 149: 1022-1031.
Reynolds W.D., Drury C.F., Tan C.S., Fox C.A. and Yang X.M. 2009. Use of indicators and pore volume function characteristics to quantify soil physical quality. Geoderma, 152: 252-263.
Safa B., Samarasinghe S. and Nejat M. 2015. Prediction of Wheat Production Using Artificial Neural Networks and Investigating Indirect Factors Affecting It: Case Study in Canterbury Province, New Zealand. Journal of Agricultural Science and Technology, 17: 791-803.
Singh Y., Singh B., Ladha J.K., Khind C.S., Gupta R.K., Meelu O.P. and Pasuquin E. 2004. Long-term effects of organic inputs on yield and soil fertility in the rice-wheat. Soil Science Society of American Journal, 68: 846-853.
Soil Survey Staff. 1996. Soil Survey Laboratory Methods Manual. Report No. 42, USDA, NRCS, NCSS, USA.

Comparison of Artificial Neural Network and Regression Models to Predict Tea Yield Based on Soil Properties

N. Yaghmaeian Mahabadi[1*], K. Samiee[2]

Assist. Prof. and M.Sc. Graduate, Dept. of Soil Science, Faculty of Agricultural Sciences, University of Guilan

Abstract

The objective of this study was to compare capability of artificial neural network (ANN) and regression models in order to predict tea yield using soil properties. The study area contained 30 hectares of tea orchard of Fouman, Guilan province. 70 soil samples were collected from 0-50 cm depth. To determine the average tea yield, green leaf sampling was conducted with the centrality of soil sampling locations. Finally, multiple linear regression and ANN models were developed for tea yield estimation in the study area and then the developed modeless were validated. The results showed that the multiple linear regression and ANN models explained 59 and 79 % of the total variability of tea yield, respectively, in the study area using soil properties. Sensitivity analysis based upon the ANN model, revealed that the percentage of clay, organic C and pH were identified as the important soil attributes influencing the tea yield distribution within study area. The overall results indicated that ANN model with ME and RMSE of 0.016 and 0.181, respectively, could provide superior predictive performance when compared with regression model.

Keywords: Yield prediction, Sensitivity analysis, Multiple linear regression