تخمین پتانسیل تولید توتون در استان مازندران با کاربرد برخی ویژگی‌های خاک

مهناز اسکندری [۱]، حامد زمانی [۲]، رضا محسن زاده [۳]
۱- محقق موسسه تحقیقات خاک و آب، ۲- محقق گروه زراعت مرکز تحقیقات و آموزش تیرتاش، ۳- محقق گروه تکنولوژی مرکز تحقیقات و آموزش تیرتاش

چکیده
تخمین مقدار تولید یک محصول زراعی با توجه به ویژگی‌های خاک، گامی ضروری در راستای برنامه‌ریزی استفاده از اراضی است. هدف از این پژوهش، بررسی کارایی شبکه‌های عصبی مصنوعی در طراحی مدلی هوشمند برای تخمین مقدار تولید توتون با توجه به برخی ویژگی‌های خاک در استان مازندران بود. به این منظور از ۴۹ مزرعه کشت توتون در این استان، یک نمونه مرکب از هر واحد جمع‌آوری و ویژگی‌هایی از خاک که در کمیت و کیفیت توتون موثر بودند، در نمونه‌ها اندازه‌گیری شدند. این ویژگی‌ها به عنوان ورودی مدل و مقدار محصول توتون بدست آمده نیز به عنوان خروجی مدل، در نظر گرفته شدند. نتایج نشان داد که یک شبکه پرسپترون چندلایه پس از آموزش قادر است مقدار تولید توتون را در مزارعی که با داده‌های آن آموزش ندیده، به خوبی تخمین زند. بنابراین مدل پیشنهادی می‌تواند تولید کمّی توتون را در یک منطقه اقلیمی مشابه با مدیریت تقریباً یکسان، با دقت مناسب برآورد نماید.

واژه‌های کلیدی: ارزیابی اراضی، پرسپترون چندلایه، شبکه‌های عصبی مصنوعی، مدلسازی

مقدمه
توتون از مشهورترین و پر مصرف‌ترین گیاهان تخدیری است که دارای مادهٔ مخدر نیکوتین بوده و با کیفیت‌های مختلف جهت مصرف به شکل‌های گوناگون مانند سیگار معمولی، پیپ، سیگار برگ، جویدن و انفیه تولید می‌شود (آهی‌فر، ۱۳۷۹). اگر چه خاستگاه توتون در نواحی حاره‌ای است، اما تولید تجاری بیشتر ارقام آن در نواحی معتدل است. تقریباً ۱۰۰ کشور تولید کنندهٔ توتون در دنیا می‌باشند که بالغ بر هفت میلیون تن محصول تولید می‌کنند. درآمد حاصل از فرآورده‌های مختلف توتون رقم مهمی از درآمد ملی کشورهای تولید کننده را تشکیل می‌دهد. طی آمار رسمی اعلام شده توسط سازمان‌های معتبر جهانی بیش از ۳/۱ میلیارد نفر در دنیا به‌طور دایمی سیگار مصرف می‌کنند که ۸۲% این افراد در کشورهای در حال توسعه زندگی می‌کنند (زمانی، ۱۳۸۹).
از مهم‌ترین نواحی کشت توتون در ایران استان‌های گلستان، مازندران و گیلان می‌باشند. در استان‌های مازندران و گلستان، توتون‌های دسته بارلی (هوا خشک)، ویرجینیا (گرمخانه‌ای) و باسما (آفتاب خشک)، در استان گیلان توتون تیپ گرمخانه‌ای، در استان کردستان توتون تیپ هوا خشک و در استان آذربایجان غربی توتون‌های تیپ هواخشک و شرقی کشت می‌گردد (آهی‌فر، ۱۳۷۹). تولید تجاری توتون در سال ۱۳۸۶ در ایران ۶۶۹۷۲۷۲ کیلوگرم بود که سطح زیر کشت آن حدود ۵۶۶۶ هکتار بوده است (بی‌نام، ۱۳۸۷). توتون در استان‌های شمالی کشور معمولاً بدون آبیاری کشت می‌شود و کیفیت آن و نحوه مصرفش تابع اثرات متقابل ژنوتیپ، خاک و عوامل اقلیمی و زراعی است.
بر اساس اعلام سازمان جهانی خواروبار و کشاورزی، سیگار و توتون از نظر ارزش، چهارمین محصول وارداتی ایران به حساب می‌آیند (زمانی، ۱۳۸۹). بنابراین توتون یک محصول مهم تجاری است که می‌تواند تولید آن در شرایط مناسب سودآور باشد. به همین دلیل تحقیقات فراوانی به منظور بررسی کمیت و کیفیت تولید آن در خاکهای مختلف انجام شده است. به عنوان مثال Zhang و همکاران (۲۰۱۵) معتقدند به دلیل وجود بیش از ۳۰۰ میلیون نفر استفاده کننده از سیگار در چین، کشت توتون از اهمیت ویژه ای برخوردار است و باید مناسب‌ترین زمینها را به کشت آن اختصاص داد. این پژوهشگران تناسب اراضی مختلف را برای کشت این گیاه در استان شاندونگ چین با استفاده از فازی، GIS و AHP بررسی کردند. در این پژوهش، ۲۰ عامل در ارزیابی در نظر گرفته شده بود. نتایج نشان داد که نزدیک به ۲۹ درصد از کل منطقه تناسب زیادی برای کشت توتون دارند. نتایج آنها روش موثری را برای افزایش کارایی استفاده از اراضی و مدیریت بهتر برای تولید توتون ارائه داد. همچنین پژوهش‌های متعددی در زمینه بررسی ارتباط وضعیت تغذیه خاک با کیفیت توتون تولید شده (مانند Zhang and Kong, ۲۰۱۴)، اثر تراکم خاک بر مقدار تولید توتون (مانند Srinivas et al., ۲۰۱۴) و تعیین مناطق مدیریتی برای کشت توتون براساس حاصلخیزی خاک (Wang et al., ۲۰۰۹) ارائه شده است.
تخمین مقدار تولید این محصول در واحدهای اراضی مختلف پیش از کشت می‌تواند گامی موثر در ارزیابی اقتصادی به شمار رود. هدف از این پژوهش، بررسی کارایی شبکه‌های عصبی مصنوعی در طراحی مدلی هوشمند برای تخمین مقدار تولید توتون با توجه به برخی ویژگی‌های خاک در استان مازندران بود.

مواد و روش‌ها

جمع‌آوری داده ها از ۴۹ مزرعه مجزای کشت توتون در استان مازندران در سال ۱۳۹۲ انجام گرفت که شرایط اقلیمی و مدیریت این مزارع تقریباً یکسان بود. برای تعیین ویژگیهای فیزیکی و شیمیایی خاک، نمونه‌هایی پیش از نشاکاری بصورت تصادفی از نقاط مختلف هر مزرعه در عمق ۰-۳۰ سانتی‌متری برداشته و از آنها یک نمونه مرکب خاک تهیه شد. نمونه ها پس از هوا خشک شدن از الک دو میلیمتری عبور داده شدند. سپس آزمایش‌های تعیین فراوانی نسبی ذرات خاک (رس، سیلت و شن)، هدایت الکتریک عصاره اشباع، اسیدیته گل اشباع، مقدار ماده آلی خاک، پتاسیم، فسفر و نیتروژن اندازه‌گیری شدند. کشت توتون گرمخانه‌ای رقم کا ۳۲۶، در این مزارع به صورت دیم انجام گرفت. مقدار توتون نیز پس از برداشت محصول در هر مزرعه، به صورت کیلوگرم در هکتار یادداشت شد.

شبکه‌های عصبی مصنوعی

در این پژوهش، از شبکه عصبی پرسپترون چند لایه (MLP)[۲۳۳] برای تخمین مقدار تولید توتون استفاده شد. شبکه‌های MLP با آموزش تحت سرپرستی و الگوریتم پس‌انتشار[۲۳۴] خطا، یکی از بهترین شبکه‌ها برای حل مسائل پیچیده و غیرخطّی، می‌باشند. در حالت کلّی، شبکه‌های MLP واجد سه لایه ورودی، پنهان و خروجی بوده که در هر لایه تعدادی واحد پردازش‌گر به نام نرون[۲۳۵] وجود دارد (نوابیان و همکاران، ۱۳۸۳). هر نرون، خروجی وزن‌دار شده از نرون قبلی را دریافت کرده و از یک تابع فعال‌سازی یا آستانه عبور می‌دهد. این توابع می‌توانند از انواع مختلف مانند سیگموئیدی، گوسی، خطی و باینری باشند. مبنای آموزش این شبکه‌ها، تغییر وزن اتصالات به منظور رسیدن به خروجی مطلوب است. بدین‌گونه که نخست الگویی به شبکه ارائه و سپس خروجی آن محاسبه می‌شود. با مقایسه این خروجی با مقدار واقعی، میزان خطا برای اصلاح وزن‌های شبکه محاسبه می‌گردد (فلامکی و همکاران، ۱۳۹۲). سپس خطا به درون شبکه انتشار یافته و وزن‌ها دوباره تنظیم می‌شوند. این چرخه تا حداقل شدن جمع مربعات خطا ادامه می‌یابد. زمانی یک شبکه را تعمیم‌یافته می‌پندارند که با کاربرد داده‌هایی که از قبل به شبکه معرفی نشده‌اند (داده‌های آزمایش) مقدار خطای پیش‌بینی قابل قبول باشد. بنابراین در مدلسازی با این شبکه‌ها از ابتدا باید داده‌ها را به دو دسته داده‌های آزمایش و آموزش تفکیک کرد. الگوی آموزش باید تا حدّ امکان کل فضای داده‌ها را پوشش دهد. برای محاسبه تعداد لایه پنهان و نرون‌های هر لایه رابطه خاصی وجود ندارد. بنابراین برای دست‌یافتن به بهترین ساختار ANN در هر مسأله مدلسازی نیاز به آزمون ساختارهای متعدد است (فلامکی و اسکندری، ۱۳۹۱).

یکی از مراحل مهم پی‌ریزی شبکه های عصبی مصنوعی، پیش پردازش داده ها است. این عمل شامل انتخاب متغیرهای موثر، انتخاب الگوهای آموزش ده و آزمون، دسته بندی الگوها و استاندارد کردن داده‌ها می‌شود. متغیرهای انتخاب شده برای تخمین مقدار تولید توتون شامل بافت خاک، درصد کربن آلی و مقدار سه عنصر غذایی ضروری برای رشد توتون یعنی نیتروژن، فسفر و پتاسیم خاک بود چراکه نشان داده شده است توتون به مقادیر عناصر ریزمغذی حساسیت کمتری دارد. همچنین مقدار شوری و اسیدیته خاک در ساخت مدل در نظر گرفته شدند. توتون به EC حساس بوده و در EC=۱ dS/m مقدار کاهش محصول ۱۰ درصدی مشاهده شده است (Sys et al., ۱۹۹۱). مقدار کلر خاک نیز به دلیل اثر بر کیفیت برگ توتون در نظر گرفته شد.

هدف از استاندارد کردن داده‌ها، هم ارزش کردن تمامی عناصر موجود در یك الگو می باشد. برای استاندارد کردن داده‌ها از رابطه زیر استفاده شد (قربانی و همکاران، ۱۳۸۸):

$$Xn = 0.1 + 0.8\left(\frac{X - Xmin}{Xmax - Xmin}\right) \quad (۱)$$

پس از انجام مراحل یادشده، داده ها به صورت تصادفی مخلوط شده و الگوهای آموزشی و آزمون به ترتیب به نسبت ۸۰ و ۲۰ درصد انتخاب گردیدند.

سنجش آماری نتایج

به منظور ارزیابی اعتبار شبکه‌های پی‌ریزی شده برای تخمین مقدار تولید توتون با توجه به برخی از ویژگی‌های خاک، از تحلیل خطاهای باقیمانده و اختلاف بین مقادیر اندازه‌گیری و پیش‌بینی شده استفاده شد. آماره‌های لازم برای این کار حداکثر خطا (ME)، میانگین ریشه دوم خطا (RMSE)، ضریب تبیین (CD)، کارآیی مدلسازی (EF)، و ضریب باقیمانده (CRM) بودند. تعریف ریاضی این آماره‌ها به صورت زیر است (Zarei et al., ۲۰۰۹):

$$ME = \max|P_i - O_i|_{i=1}^{n} \quad (۲)$$

[۲۳۳] Multi Layer Perceptrone
[۲۳۴] Back propagation
[۲۳۵] Neuron

$$RMSE=\left[\frac{\sum_{i=1}^{n}\left(P_i-O_i\right)^2}{n}\right]^{0.5}\frac{100}{\bar{O}} \quad (۳)$$

$$CD=\frac{\sum_{i=1}^{n}\left(O_i-\bar{O}\right)^2}{\sum_{i=1}^{n}\left(p_i-\bar{O}\right)^2} \quad (۴)$$

$$EF=\frac{\sum_{i=1}^{n}\left(O_i-\bar{O}\right)^2-\sum_{i=1}^{n}\left(P_i-O_i\right)^2}{\sum_{i=1}^{n}\left(O_i-\bar{O}\right)^2} \quad (۵)$$

$$CRM=\frac{\sum_{i=1}^{n}O_i-\sum_{i=1}^{n}P_i}{\sum_{i=1}^{n}O_i} \quad (۶)$$

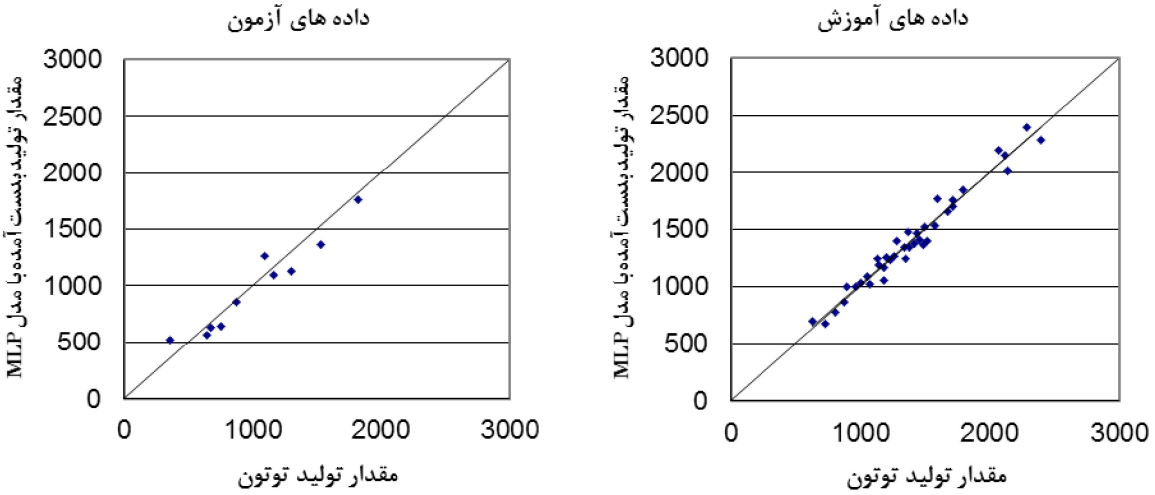
که در آن‌ها، Pi مقدار پیش‌بینی شده، Oi مقدار اندازه‌گیری شده (مشاهده شده)، n تعداد نمونه‌های به کار رفته و $\bar{O}$ مقدار متوسط مقادیر O است. چنانچه تمام مقادیر اندازه‌گیری شده و برآوردشده با هم برابر باشند، مقدار عددی این آماره‌ها برابر است با: EF=۰، CD=۱، RMSE=۰، ME=۰ و CRM=۰.

نتایج و بحث

برای دستیابی به بهترین مدل تخمین مقدار تولید توتون در مزارع استان مازندران، تعداد زیادی شبکه عصبی MLP با ساختار متفاوت ارزیابی شدند. نتایج حاصل از بهترین مدلهای طراحی شده در جدول ۱ نشان داده شده است. این جدول نشان می‌دهد که مدل طراحی شده بر پایه شبکه عصبی، می تواند به خوبی با داده‌های ورودی آموزش داده شده و مقدار خروجی را تخمین بزند. در این میان، بهترین مدل برازش داده شده به داده‌های آموزش و آزمون، شبکه‌ای با ساختار ۱-۷-۸ (به ترتیب تعداد نرون در لایه خروجی، پنهان و ورودی) با تابع محرک تانژانت سیگموئید بود. مقایسه مقدار توتون ماحصل از مزارع در سال ۱۳۹۲ با مقدار خروجی تخمین زده شده به کمک این شبکه با کاربرد داده‌های آموزش و آزمون، در شکل یک نشان داده شده است.

جدول ۱- ویژگی‌های مناسب‌ترین شبکه‌های عصبی طرح‌ریزی شده در مدلسازی تخمین مقدار توتون

نوع شبکه	ساختار	داده‌های آموزش		داده‌های آزمون	
		ضریب همبستگی	RMSE	ضریب همبستگی	RMSE
MLP	۱-۳-۸	۸۹/۰	۰۴/۰	۸۲/۰	۲۲/۱
MLP	۱-۳۰-۸	۸۱/۰	۰۶/۰	۷۶/۰	۴۰/۲
MLP	۱-۱۵-۸	۹۲/۰	۰۳/۰	۸۷/۰	۳۵/۱
MLP	۱-۳-۲-۸	۹۳/۰	۰۳/۰	۸۶/۰	۶۷/۱
MLP	۱-۵-۳-۸	۸۹/۰	۰۵/۰	۸۰/۰	۳۲/۱
MLP	۱-۵-۸	۹۰/۰	۰۳/۰	۸۷/۰	۸۹/۳
MLP	۱-۷-۸	۹۶/۰	۰۲/۰	۹۳/۰	۸۷/۰



شکل ۱ - مقایسه مقدار تولید بدست آمده از مزارع توتون با مقدار تخمینی به کمک مدل شبکه عصبی

خط تناظر یک به یک نیز در این شکل نمایان است. هرچه پراکنش داده‌ها در اطراف این خط کمتر باشد، نشان می‌دهد که مدل تخمین مقادیر خروجی بهتر عمل کرده است. ضریب هم‌بستگی میان تولید واقعی توتون و مقدار تخمینی در این مدل، ۹۶/۰ برای داده‌های آزمونی و ۹۳/۰ برای داده‌های آزمون بدست آمد.

آماره‌های سنجش این مدل برای هر ۴۹ مجموعه داده‌های کاربردی، در جدول ۲ ارائه شده است. آماره CRM نشان می‌دهد که مدل شبکه عصبی اندکی کم‌برآورد عمل کرده است و مقدار تخمینی را کمی کمتر از واقعیت نشان می‌دهد. هر چقدر آماره EF به یک نزدیک‌تر باشد، نشان دهنده کارایی بهتر مدل کاربردی است. بنابراین بر پایه نتایج بدست آمده می‌توان اینگونه پنداشت که مدل شبکه‌های عصبی پرسپترون چندلایه طراحی شده، به خوبی قادر است در شرایط اقلیمی و مدیریتی یکسان، مقدار تولید توتون را به کمک برخی ویژگی‌های خاک تخمین زند. این مهم نشان می‌دهد که امکان استفاده از این مدل برای سایر محصولات زراعی نیز وجود دارد که می‌تواند مدلی دقیق در تخمین مقدار محصول پیش از کاشت گیاه در هر واحد اراضی باشد. بدین ترتیب در هر منطقه اقلیمی با مشارکت این مدل در انتخاب تیپ‌های بهره‌وری از اراضی که از نظر اقتصادی سودآور هستند، می‌توان الگوی کشت مناسب را تعیین کرد.

جدول ۲- پارامترهای آماری ارزیابی مدل شبکه عصبی

مدل‌ها	آماره‌ها				
	ME	RMSE	CD	EF	CRM
۸-۷-۱	۱۷۰	۴۴/۰	۹۴/۰	۹۷/۰	۰۲/۰

منابع

آهی‌فر، ح. ۱۳۷۹. بوتانیک و مشخصات توتونهای ایران. شرکت دخانیات ایران.

بی‌نام. ۱۳۸۷. کارنامه آماری شرکت دخانیات ایران.

زمانی، پ. ۱۳۸۹. زراعت و عمل‌آوری توتون. چاپ اول.

فلامکی، ا. اسکندری، م. بغلانی، ع.ا. و احمدی، س.ا. ۱۳۹۲. مدلسازی بار رسوب کل رودخانه‌ها با استفاده از شبکه‌های عصبی مصنوعی. حفاظت منابع آب و خاک، ۲ (۳): ۱۳-۲۵.

فلامکی، ا. و اسکندری، م. ۱۳۹۱. تخمین ضریب توزیع خاک-آب فلزات سنگین با کاربرد شبکه‌های عصبی مصنوعی. حفاظت منابع آب و خاک، ۲ (۱): ۳۶-۲۵.

قربانی دشتکی، ش. همایی، م. و مهدیان، م.ح. ۱۳۸۸. برآورد نفوذ آب به خاک با استفاده از شبکه‌های عصبی مصنوعی. مجله آب و خاک، ۲۳(۱): ۱۸۵-۱۹۸.

نوابیان، م. لیاقت، ع. و همایی، م. ۱۳۸۳. برآورد سریع هدایت آبی اشباع خاک با استفاده از شبکه‌های عصبی مصنوعی. دومین کنفرانس ملی دانشجویی منابع آب و خاک، دانشکده کشاورزی، دانشگاه شیراز.

Srinivas P. Ramakkrushnan S. Vijayan A. ۲۰۱۴. A Study on Soil Compaction Management in Tobacco Cultivation in Mysore Region of India. APCBEE Procedia, ۸: ۲۸۷-۲۹۲.

Sys C. Vanranst E. and Debaveye J. ۱۹۹۱. Land evaluation, part III. International training center for post graduate soil scientist, Ghent University, Ghent.

Wang X-Z. Liu G-S. Hu H-C. Wang Z-H. Liu Q-H. Liu X-F. Hao W-H. Li Y-T. ۲۰۰۹. Determination of management zones for a tobacco field based on soil fertility. Computers and Electronics in Agriculture, ۶۵(۲): ۱۶۸-۱۷۵.

Zarei G., Homaee M. and Liaghat A. ۲۰۰۹. Modeling transient evaporation from descending shallow groundwater table based on Brooks-Corey retention function. Water Resources Management, ۲۳(۱۴): ۲۸۶۷-۲۸۷۶.

Zhang C. and Konga F. ۲۰۱۴. Isolation and identification of potassium-solubilizing bacteria from tobacco rhizospheric soil and their effect on tobacco plants. Applied Soil Ecology, ۸۲: ۱۸-۲۵.

Zhang J., Su Y., Wu J. and Liang H. ۲۰۱۵. GIS based land suitability assessment for tobacco production using AHP and fuzzy set in Shandong province of China. Computers and Electronics in Agriculture, ۱۱۴: ۲۰۲-۲۱۱.

Abstract

Estimation of crop yield considering soil characteristics is an important stage in land use planning. This study was aimed to evaluate the effectiveness of artificial neural networks in the intelligent design model to estimate the tobacco yield in Mazandaran province. Composite soil samples were collected from each ۴۹ tobacco farms in this province. Effective soil properties in the quantity and quality of tobacco and the actual tobacco yield were considered as model input and output, respectively. The results showed that the proposed model can accurately estimate tobacco yield in the similar climate zones with almost identical management.