نشریه پژوهش‌های زعفران (دو فصلنامه)
جلد نهم، شماره دوم، پاییز و زمستان ۱۴۰۰
شماره صفحه: ۳۶۷-۳۵۲
http://dx.doi.org/10.22077/JSR.2021.4399.1162

مقاله پژوهشی:

پیش‌بینی مقدار گل و کلاله زعفران براساس خصوصیات فیزیکی و شیمیایی آب و خاک با استفاده از مدل‌های رگرسیونی چند متغیره خطی و درخت تصمیم M5

مرتضی ریوندی[۱]، عظیم قاسم‌نژاد[۲]*، خلیل قربانی[۳]،خدایار همتی[۴]، عباس ابهری[۵]
۱- دانشجوی کارشناسی ارشد دانشگاه منابع طبیعی و کشاورزی گرگان، ایران.
۲، ۳- دانشیار گروه باغبانی دانشگاه منابع طبیعی و کشاورزی گرگان، ایران.
۴- دانشیار گروه مهندسی آب دانشگاه منابع طبیعی و کشاورزی گرگان، ایران.
۵- استادیار گروه کشاورزی دانشگاه پیام نور، ۴۶۹۷-۱۹۳۹۵ تهران، ایران.
*نویسنده مسئول *Email: ghasemnezhad@gau.ac.ir*

تاریخ دریافت: ۱۴۰۰/۰۳/۱۲؛ تاریخ پذیرش: ۱۴۰۰/۰۶/۰۸

چکیده

رشد مهیج علوم وفنون مختلف و پیچیده‌تر شدن تصمیم‌گیری‌ها در دهه‌های اخیر، سر فصل‌های تازه‌ای را برای بشریت رقم زده تا با استفاده از سیستم‌های اطلاعاتی و هوش مصنوعی، با دقت و سرعت بیشتری کارهای خود را انجام داده و برای پیش‌بینی و محاسبات وقت‌گیر علمی و فنی خود راه‌کاری ارائه دهد. این تحقیق به منظور ارزیابی برآیند پیش‌بینی دو مدل رگرسیون گام به گام و مدل درخت تصمیم M5 تحت تاثیر خصوصیات مختلف آب و خاک بر عملکرد گل و کلاله زعفران در سال۱۳۹۷ در مزارع زعفران‌کاری سبزوار (واقع در طول جغرافیایی "۵۷/۴۳" عرض جغرافیایی "۳۶/۱۲") و در آزمایشگاه دانشکده تولیدگیاهی دانشگاه منابع طبیعی و کشاورزی گرگان انجام شد. در فروردین ماه پس از پایان فصل رشد زعفران از ۶۹ مزرعه زعفران‌کاری شده نمونه خاک از عمق صفر تا ۳۰ سانتی‌متری و ۱۲ نمونه آب آبیاری این مزارع تهیه و جمع‌آوری شده و جهت انجام آزمایشات خاک و اندازه‌گیری خصوصیات فیزیک و شیمیایی نمونه‌های خاک، ۱۳ پارامتر از جمله pH، اسیدیته، درصد اجزاء خاک، عناصر خاک و .... همچنین برخی از پارامترهای آب، ۴ پارامتر، مانند اسیدیته، بی‌کربنات و ....، به آزمایشگاه منتقل شد. گل‌ها در زمان ظهور گل از سطح مناطق مشخص شده مزارع جمع‌آوری شده و اندازه‌گیری‌های مورد نظر انجام شد. نتایج نشان داد با توجه به صرف وقت و هزینه-های بالای آزمایشات آب و خاک، مدل درخت تصمیم M5، از دقت و سرعت بیشتری و هزینه کمتری نسبت به مدل رگرسیون، برخوردار است. بطوری که در برآیند پیش‌بینی مدل رگرسیونی گام به گام، در ایده‌آل‌ترین حالت و ورود تمامی پارامترهای اندازه گیری شده، وزن کلاله خشک و وزن گل به ترتیب با همبستگی‌های ۷۰ و ۷۴ درصد و مقدار خطا برابر RMSE=۰/۲۳ و RMSE=۱۶/۳۸ پیش‌بینی شد. در حالی که مدل درخت تصمیم M5 با وارد کردن پارامترهای کمتری از توانمندی بالایی در جهت پیش‌بینی وزن گل و وزن کلاله خشک برخوردار بود. به‌طوری‌که وزن کلاله خشک و وزن گل را با ۹۰ درصد همبستگی و مقدار خطای برابر با RMSE=۰/۱۲ و RMSE=۹/۴ در انتهای مدل‌سازی، برای منطقه مورد مطالعه برآورد کرد. بنابراین، روش درخت تصمیم M5 در ارزیابی و پیش‌بینی عوامل مختلف بر عملکرد زعفران توصیه می‌شود.

واژه‌های کلیدی: مدل‌سازی، هوش مصنوعی، عملکرد زعفران، عناصر خاک.

مقدمه

رشد مهیج علوم و فنون مختلف و پیچیده‌تر شدن تصمیم-گیری‌ها در دهه‌های اخیر، انسان‌ها را واداشته تا با استفاده از سیستم‌های اطلاعاتی و هوش مصنوعی، با دقت و سرعت بیشتری کارهای خود را انجام دهند. شاید تولید ربات‌ها به عنوان اولین ماشین‌های دارای هوش مصنوعی که توانایی انجام کارهای دقیق و حساس را دارند سر فصل‌های تازه‌ای را برای بشریت رقم زد تا بتواند برای پیش بینی و محاسبات وقت‌گیر علمی و فنی خود راه‌کاری ارائه دهد. استفاده از هوش مصنوعی و ابزارهای داده کاوی در کشاورزی که کلیدی‌ترین مسئله بقای انسان است، توجه دانشمندان این عرصه را به خود جلب کرده‌است (*Dousti Irani, Golzarian, 2016*).

سیستم‌هایی هوشمند در واقع مانند مغز انسان مسایل را تجزیه و تحلیل می‌کنند و با استفاده از روش‌های آماری موجود بهترین مدل عملکردی را ارائه می‌دهند. امروزه از تکنیک‌های رگرسیونی، درخت‌های تصمیم، شبکه‌های عصبی مصنوعی و... در علوم مختلف و در جهت پیش‌بینی با دقت بالا استفاده می‌شود. این تکنیک‌ها رابطه‌ای را بین داده‌های ورودی با داده‌های خروجی پیش‌بینی می‌کنند. در تحقیق پیشرو از گیاه دارویی زعفران که دارای ارزش اقتصادی زیادی است و در حال حاضر در مناطق خشک و نیمه‌خشک، بویژه خراسان‌های رضوی و جنوبی کشت می-شود، استفاده شد و نهایتا توان عملکردی این گیاه در آب و خاک‌های نامرغوب و شور، سنجیده شد و مدل سازی قابل پیش بینی ارئه گردید.

خصوصیات خاک و آب نقش مهمی در تعیین اثر متقابل گیاه با خاک، جذب آب و مواد غذایی، توسعه سیستم ریشه و در نهایت رشد گیاه دارند (*Gresta et al., 2008, Maleki et al., 2019*) در یک بررسی رضوانی مقدم و همکاران (*Rezvani Moghaddam et al., 2015*) نشان داده‌اند که تمام پارامترهای فیزیکو شیمیایی خاک مورد آزمایش بر رشد بنه و عملکرد زعفران موثر بود (*Rezvani Moghaddam et al., 2015*).

تکنیک داده‌کاوی به‌عنوان یکی از محدود تکنولوژی‌های برتر که تحول چشم‌گیری در تجزیه و تحلیل اطلاعات انجام خواهد داد، مورد استفاده می‌باشد. در این تکنیک داده‌ها به دو صورت تجزیه و تحلیل می‌شود یا توصیفی هستند که خواص عمومی داده‌ها را تفسیر کرده و با نمایش الگویی مناسب از داده‌ها، تفسیر قابل درکی به انسان می‌دهند و یا با بکارگیری چند متغیر، پیش‌بینی مقادیر آینده یا مقادیر ناشناخته را پیشگویی می‌کند (*Hosseini et al., 2013*).

امروزه کاربردهای آمار مخصوصاً آمار پیشرفته در مطالعات کشاورزی روز به روز گسترش قابل توجهی پیدا کرده و از اهمیت زیادی برخوردار است. از فنون بسیار مهم در مطالعات کشاورزی انجام تحلیل‌های رگرسیونی است. در تحلیل‌های رگرسیونی توجه به این مسئله که چگونه متغیرهای مستقل، مقادیر متغیرهای وابسته را تحت تأثیرخود قرار می دهند، دارای اهمیت ویژه است. اینکه چه متغیرهایی را مستقل وچه متغیرهایی را وابسته در نظر گرفته شود، همیشه به طور کامل روشن نمی باشد و وابسته به اهداف تحقیق است (*Kalantari, 2008*). زمانی یک مدل رگرسیونی قابلیت اجرا دارد که، با توجه به تحلیل‌های رگرسیونی، مناسب مدل شدن باشد (*Salehi Sadeghian and Ebrahimi, 2002*).

یکی دیگر از مشهورترین روش‌ها در داده‌کاوی، استفاده از درخت تصمیم[1] می‌باشد. مدل‌های درخت جهت پیش‌بینی مقدار تاثیر گذاری عوامل دخیل در یک فرایند مورد استفاده قرار می‌گیرند. سادگی و انعطاف پذیری این روش نسبت به سایر روش‌های داده کاوی، علت اصلی گرایش به آن است (*Cichosz ،2014*).

ویژگی‌های درخت تصمیم

درخت تصمیم دارای ویژه‌گیهای مهمی است. از جمله: ۱- داده‌هایی را که حجم بالایی دارند به راحتی آنالیز می‌شود و از این رو در داده کاوی بسیارمورد توجه می-باشد، ۲- در رابطه با داده‌های گمشده و داده‌های ورودی قدرتمندانه عمل می‌کند، ۳- با استفاده از قوانین *if-then* نتایج، قابل فهم‌تر برای استفاده است،۴- امکان ترکیب عطفی و فصلی فرضیه‌ها را می‌دهد، ۵- رویه آموزش مدل ساده بوده و نیاز به آزمون وخطا ندارد (*Ghorbani et al., 2016*).

در بسیاری از پژوهش‌ها و مطالعات کشاورزی، همواره برآیند چگونگی تاثیرگذاری متغیرها بر یکدیگر، کنجکاوی محققان را برانگیخته است. از این رو روش-های به روز شده‌ی علمی متفاوتی را همواره سرلوحه کار و تحقیق خود قرار داده‌اند. تا ضمن رسیدن به نتایج قابل قبول‌تر، از سنگینی تجزیه و تحلیل‌های آماری

[1] Decision tree model

دست و پا گیر رهایی یابند. طراحی مدل‌های گوناگون و پیش‌بینی مناسب نتایج تحقیقات، آرزویی است که داده کاوی‌های جدید آن را تا حدودی محقق کرده‌است.

با پژوهشی که، ریمسان و همکاران (*Remesan et al., 2008*) در حوضه‌ی آبریز برو[2] انگلستان انجام دادند، عوامل موثر برتابش خورشیدی را با استفاده از دو مدل رگرسیونی و شبکه عصبی پیش‌بینی کردند. نتایج بدست آمده، دلالت بر دقت بیشتر مدل رگرسیونی داشت.

همچنین مطالعاتی که، لئو و همکاران (*Luo et al., 2008*)، در پیش بینی کربن آلی و محمودآبادی و کریمی (*Mahmoudabadi & Karimi, 2015*)، در تغییرات رس بافت خاک با استفاده از مدل‌های رگرسیونی انجام دادند به کارا بودن مدل‌های رگرسیونی در داده کاوی اذعان داشتند.

در تحقیقی که ستاری و همکاران (*Sattari et al., 2013*) بر روی رودخانه سوهای آنکارا انجام دادن، با استفاده از توانمندی تکنیک درخت *M5*، جریان آب این رودخانه را پیش‌بینی کردند. در مطالعه مدل سازی تشت تبخیر که ترزی (*Terzi 2007*) با استفاده از مدل‌های داده کاوی، شبکه عصبی، درخت *M5*، *M5Rules*، *Kstar* و روش‌های رگرسیونی بر روی دریاچه آگیردیر ترکیه انجام داد، ضمن تایید قدرتمند بودن تکنیک *M5* نسبت به سایر روش‌های مدل سازی شده، موثرترین داده‌های هواشناسی را بر روی تشت تبخیر، پارامترهای رطوبت نسبی، دمای آب و هوا و انرژی تابشی خورشید دانست. نتایج مشابه ای از بررسی‌های قبلی مولف و همکاران در این خصوص گزارش شد (*Rivandi et al,. 2020*). از منابع استفاده شده چنین برداشت می‌گردد که بیشتر مدل سازی‌های انجام شده در حوزه‌های گوناگون کشاورزی، در حوضه-های آب، خاک و هواشناسی بوده و در زمینه عملکرد محصولات کشاورزی و خصوصا گیاهان دارویی تحقیقات چندانی صورت نگرفته است. لذا مطالعه حاضر با هدف بررسی روابط خطی و غیر خطی بین خصوصیات فیزیک و شیمیایی آب و خاک با مقدار گل و کلاله زعفران بر اساس مدلهای رگرسیونی چند متغیره خطی و درخت تصمیم *M5* می باشد.

مواد و روش‌ها

به منظور بررسی ارزیابی دو مدل رگرسیون گام به گام و درخت تصمیم *M5* در برآیند پیش‌بینی گل و کلاله زعفران، که با استفاده از خصوصیات مختلف آب و خاک دخیل در عملکرد زعفران در سال۱۳۹۷ در مزارع زعفران‌کاری سبزوار انجام گرفت. سبزوار یکی از شهرستان‌های استان خراسان رضوی بوده و فاصله آن تا مشهد ۲۴۰ کیلو متر است. ارتفاع از سطح دریا ۹۷۷ متر، طول جغرافیایی "۵۷،۴۳" عرض جغرافیایی"۳۶،۱۲" می‌باشد. دارای آب و هوای بیابانی با تابستان‌های گرم و زمستان‌های سرد و خشک است.

در شروع کار پس از پایان فصل رشد زعفران که مصادف با اواخر فروردین‌ماه و اوایل اردیبهشت‌ماه بود، از ۶۹ مزرعه زعفران و از هر مزرعه بصورت زیکزاکی پنج نمونه خاک از عمق صفر تا ۳۰ سانتی‌متری خاک جمع‌آوری شد که پس از مخلوط کردن خاک هر مزرعه مقداری به عنوان نمونه جهت کارهای آزمایشگاهی کدگذاری و پس از خشک‌کردن نمونه‌های خاک، جهت انجام آزمایشات خاک‌شناسی به آزمایشگاه باغبانی، دانشکده تولید گیاهی منتقل شد. پارامترهای اندازه-گیری شده از نمونه خاک‌ها در سه تکرار که شامل اندازه‌گیری ماده (کربن) آلی (*Nelson, 1982*)، تعیین بافت خاک به روش هیدرومتری (*Movahedi Naeini & Rezaei, 2008*)، اندازه‌گیری اسیدیته خاک (*pH*) و هدایت الکتریکی خاک (*EC*) (*Pich et al., 1982*)، اندازه‌گیری فسفر قابل جذب (*Olsen.,1954*)، اندازه‌گیری پتاسیم و سدیم قابل جذب به روش شعله‌سنجی (*Pich et al., 1992*)، اندازه‌گیری میزان ازت کل با کجلدال (*Bremner & Mulvaney1982*)، اندازه‌گیری کلسیم، منیزیم و کلر(*Pich et al., 1982*) بود (جدول ۱).

جمع‌آوری آب آبیاری

جمع‌آوری آب مورد آبیاری مزارع زعفران همزمان با اولین آبیاری مزرعه انجام شد. نمونه‌ها در سه تکرار جهت اندازه‌گیری خصوصیات شیمیایی شامل اندازه‌گیری اسیدیته آب (*pH*) و هدایت الکتریکی آب (*EC*) (*Pich et al., 1982*) و اندازه‌گیری کربنات و بی‌کربنات آب (*Richard., 1954*) به آزمایشگاه منتقل شد (جدول ۱).

[2] Brue

جمع‌آوری و اندازه‌گیری گل و کلاله زعفران

قبل از آغاز گلدهی زعفران در مزارع آزمایشی، یک متر مربع از هر زمین کادر بندی شد و پس از شروع دوره‌ی گلدهی که در اواسط آبان ماه ۱۳۹۷ شروع شد و تا ۱۰ آذرماه ادامه داشت، همزمان جمع‌آوری، خشک کردن و اندازه‌گیری گل و کلاله‌ی زعفران، انجام شد (جدول ۱).

جدول ۱- توصیف آماری خصوصیات فیزیک و شیمیایی آب و خاک مزارع مورد مطالعه

Table 1 - Statistical description of physical and chemical properties of water and soil of the studied farms

متغیرها	واحد	میانگین	بیشینه	کمینه	دامنه تغیرات	انحراف معیار	واریانس	RMS
pH خاک	-	8.08	8.51	7.28	1.23	0.265	0.06	85.0
EC خاک	dS/m	5724.38	15750	1024	14726	2683.8	7202786	6314.02
EC آب	dS/m	5641.86	9020	1409	7611	2080.8	4329745	6008.12
pH آب	-	7.583	8.3	7.06	1.27	0.232	0.054	7.587
HCO_3- آب	meq/lit	21.3	60	10	50	9.844	96.92	23.439
عمق کاشت	cm	17.36	27.5	12	15.5	2.495	6.225	17.545
سن	year	3.76	10	2	8	1.582	2.504	4.082
شن	%	50.85	88.4	32.24	56.16	10.587	112.089	51.929
رس	%	18.25	34.04	5.32	28.72	4.846	23.49	18.875
سیلت	%	30.84	50	6.28	43.72	8.001	64.016	31.855
ماده آلی	%	1.83	2.77	0.94	0.94	0.375	0.141	1.877
سدیم	meq/lit	40.43	199.54	14.18	185.36	26.898	723.544	48.497
پتاسیم	meq/lit	53.48	225.88	26.08	199.8	32.165	1034.59	62.295
کلر	meq/lit	48.53	147.5	5	142.5	28.545	814.841	56.2
کلسیم	meq/lit	7.04	23.6	0	23.6	5.173	26.762	8.721
منیزیم	meq/lit	30.96	85.2	2.4	82.8	16.004	256.143	34.805
فسفر	meq/lit	18.19	46.14	4.64	41.5	8.585	73.708	20.092
نیتروژن	meq/lit	0.04	0.3	0	0.308	0.051	0.0027	0.069
وزن گل	gr/m^2	18.29	100.42	0.96	99.46	22.294	497.031	28.715
وزن کلاله خشک	gr/m^2	0.27	1.308	0.018	1.29	0.296	0.087	0.401

رگرسیون

تحلیل تاثیرات بوجود آمده بین دو متغییر، یکی به عنوان متغییر مستقل و دیگری به عنوان متغییر وابسته، جهت برآورد عمل پیش بینی از یک پدیده را، رگرسیون می‌گویند. انواع رگرسیون بصورت دو متغیره و چند متغیره محاسبه می‌شود. در رگرسیون دو متغیره یک متغیر مستقل و یک متغییر وابسته وجود دارد. مثلا در محاسبه تاثیر دو متغیر شوری خاک و وزن خشک کلاله زعفران، شوری متغیری است مستقل و وزن خشک کلاله متغییر وابسته‌ای است که به مقدار شوری خاک وابسته است. در رگرسیون چند متغییره همواره دو یا بیش از دو متغییر مستقل وجود دارد که بر روی یک متغییر وابسته تاثیرگذار است. بطور مثال، چگونگی تاثیرات متغییرهای مستقلی مانند، مقدار ماده آلی خاک، اسیدیته آب و سدیم خاک بر متغییر وابسته‌ای مثل، وزن پیاز زعفران که متاثر از این پارامترها است، نمایش داده می‌شود. مدل رگرسیونی اگر دارای یک متغیر مستقل باشد، از (رابطه ۱) و اگر بیشتر از یک متغیر مستقل داشته‌باشد از (رابطه ۲) و چند متغیره نامیده می‌شود.

رابطه

$$Y=\beta0+ \beta1X1 + e \quad (۱)$$

رابطه

$$Y= \beta0 + \beta1X1 + \beta2X2 + \ldots +\beta nXn + e \quad (۲)$$

(Y): متغییر وابسته، X1....Xn : متغییرهای مستقل، β0…βn: پارامترهای مدل هستند (مونت گومری و همکاران، ۲۰۱۵).

مدل درخت M5

در سال ۱۹۹۲ میلادی کوینلان (Quinlan) از مدلی درخت به نام M5 رو نمایی کرد که امروزه در داده کاوی بسیار مورد توجه است. این مدل یک درخت تصمیم باینری[3]یا دودویی است که در هر گره انتهایی با روابط رگرسیونی خطی مقدار عددی را پیش‌بینی می-نماید. انجام مدل درخت در دو مرحله صورت می‌گیرد: مرحله اول: گسترش و ساخت درخت تا زمانی که به برگ‌های همگن و خالص برسد. مرحله دوم: هرس شاخه‌هایی که دقت مدل را کاهش می‌دهند (*Fallahi et al., 2011*).

آنالیز داده‌ها

پس از جمع‌آوری داده‌های خام و دسته بندی‌های لازم، کلیه جداول و نمودارها توسط نرم‌افزار Excel 2013 رسم شد.

این داده‌ها با استفاده از نرم‌افزارهای WEKA[4] که توسعه آن در دانشگاه Waikato نیوزلند بوده و مینی‌تب[5] روش‌های آماری معمول مثل رگرسیون و درخت تصمیم M5، آنالیز شدند.

نتایج وبحث

تحلیل رگرسیون گام به گام در بررسی برآورد وزن گل اندازه‌گیری شده

رگرسیون چند متغیره خطی به روش گام به گام اجراء شد و در هر گام بهترین زیرمجموعه یک تا چند عضوی که بالاترین میزان همبستگی را ایجاد کردند، انتخاب شدند (جدول ۲). در گام اول مقدار EC آب به تنهایی ۶۰/۸ درصد همبستگی را با وزن گل برقرار می‌کند. وارد کردن متغیر سیلت به مدل باعث بهبود نتایج می‌شود و درصد همبستگی را تا ۶۳/۷ درصد افزایش می‌دهد. مقدار بی‌کربنات و ماده آلی نیز در گام‌های بعدی به ترتیب باعث بهبود درصد همبستگی تا ۶۵/۴ و ۶۷/۱ درصد می‌شوند (جدول ۲ و شکل ۱). افزودن متغیرهای بیشتر باعث می‌شود تغییر زیادی در درصد همبستگی ایجاد نشود بطوری که با دخالت دادن تمام ۱۸ متغیر اندازه‌گیری شده در مدل به عنوان متغیرهای مستقل یا پیش‌بین درصد همبستگی درنهایت به ۷۴ درصد افزایش خواهد یافت اما مدلی با ۱۸ متغیر به مراتب خیلی پیچیده‌تر از مدلی با ۴ متغیر خواهد بود و از عدم قطعیت بالاتری نیز برخوردار خواهد بود.

تحلیل رگرسیون گام به گام در بررسی برآورد وزن کلاله خشک

در مورد وزن کلاله نیز، مدل رگرسیون چند متغیره خطی به شکل گام به گام اجراء شد تا نقش متغیرهای پیش‌بین یا مستقل در مدل دیده شود. نتایج نشان داد در این مورد نیز میزان EC آب بالاترین همبستگی معادل ۵۴/۸ درصد را با وزن کلاله در مقایسه با دیگر متغیرها ایجاد می‌کند. در گام‌های بعدی درصد شن، PHآب و درصد پتاسیم باعث می‌شود همبستگی تا ۶۲/۹ درصد افزایش یابد (جدول ۳). افزودن متغیرهای دیگر به مجموعه متغیرهای مستقل، در نهایت درصد همبستگی را تا ۷۰ درصد افزایش می‌دهد. از اینرو مدلی با ۴ متغیر به دلیل سادگی و پیچیدگی کمتر به عنوان مدل مناسب‌تر انتخاب شد (شکل ۲).

نمودار پراکنش مقادیر پیش‌بینی شده در مقابل مقادیر اندازه‌گیری شده (شکل ۲) نشان می‌دهد مدل رگرسیونی چند متغیره علاوه بر دقت پایین در برآورد وزن کلاله خشک، احتمالا قادر به برآورد عملکردهای بالای ۰/۷ گرم بر متر مربع نیز نمی‌باشد. مگر با اضافه شدن پارامترهای دیگری مانند پارامترهای اقلیمی، شیب زمین، ارتفاع از سطح دریا و .....، که نیازمند تحقیقات بیشتری می‌باشد.

تحلیل درخت تصمیم M5 در برآورد وزن کلاله خشک

با توجه به نتایج نه چندان رضایت بخش مدل رگرسیون چند متغیره خطی، از مدل درخت تصمیم M5 استفاده شد تا روابط غیرخطی بین داده‌ها را با تفکیک داده‌ها به چندین زیرمجموعه و ارائه روابط ساده رگرسیونی چند متغیره خطی برای هر یک از آنها مدل کند. در مورد وزن کلاله خشک، فضای داده‌ها به ۹ زیرمجموعه تفکیک شد. در ابتدا مجموعه داده‌ها بر اساس سدیم قابل جذب خاک به دو دسته تقسیم‌بندی شدند.

[3] Binary
[4] Waikato Environment for Knowledge Analysis
[5] Minitab

جدول ۲- آنالیز رگرسیون گام به گام در بررسی وزن گل اندازه‌گیری شده

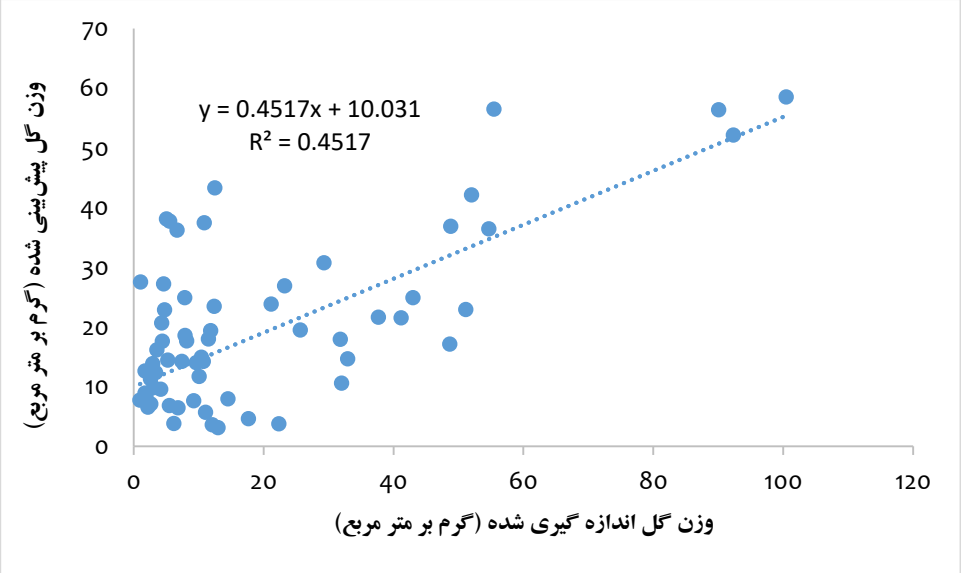
Table 2- Stepwise regression analysis investigating the measured flower weight

	EC آب	سیلت	بی‌کربنات	ماده آلی	درصد همبستگی	R-Sq
گام اول	-0.00652	-	-	-	60.8	٪37
گام دوم	-0.00506	0.660	-	-	63.7	٪40.75
گام سوم	-0.00527	0.577	0.337	-	65.4	٪42.9
گام چهارم	-0.00543	0.745	0.456	10.47	67.1	٪45.17

رابطه ۱- رابطه رگرسیونی چند متغیره خطی جهت برآورد وزن گل:

$$\text{وزن گل} = 35.5 - 0.00543\ W\ EC + 0.456\ CO3^{-} + 0.745\ silt - 10.47\ MO$$

هدایت الکتریکی آب: W EC، بی کربنات آب: CO3، سیلت: silt، ماده آلی: MO



شکل ۱- نمودار پراکنش مقادیر وزن گل برآورد شده بر اساس مدل رگرسیون چند متغیره خطی و مقادیر مشاهداتی

Figure 1- Distribution diagram of flower weight values estimated based on multivariate linear regression model and observational values

جدول ۳- آنالیز رگرسیون گام به گام در بررسی وزن کلاله خشک اندازه‌گیری شده

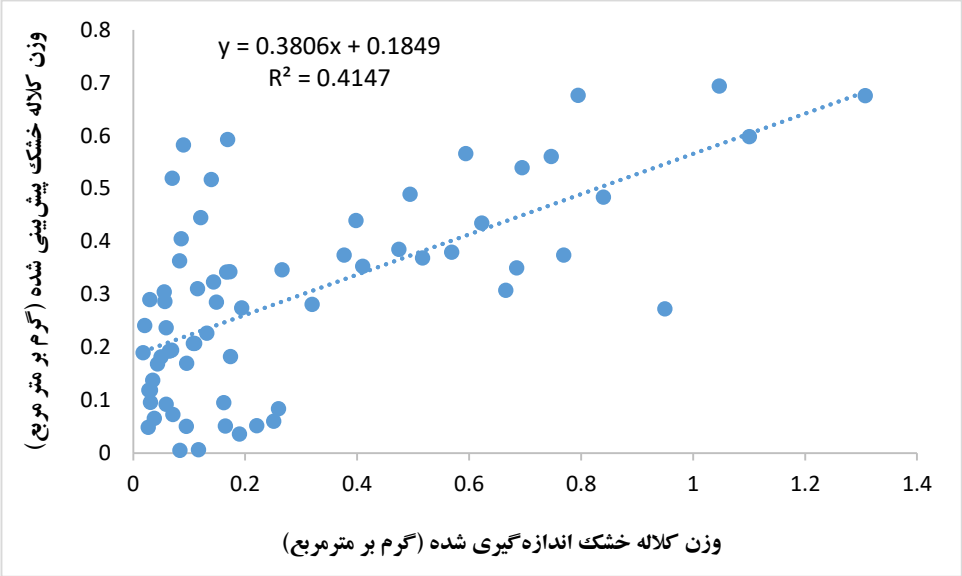
Table 3- Stepwise regression analysis in the study of measured dry stigma weight

	همبستگی (درصد)	پتاسیم	pH آب	شن	EC آب	R-Sq (درصد)
گام اول	54.8	-	-	-	-0.000081	30.12٪
گام دوم	57.6	-	-	-0.00571	-0.000068	٪33.20
گام سوم	61.2	-	-0.408	-0.00571	-0.000096	٪37.50
گام چهارم	62.9	-0.001450	-0.388	-0.00779	-0.000099	٪39.67

رابطه ۲- رابطه رگرسیونی چند متغیره خطی جهت برآورد وزن کلاله خشک:

$$\text{کلاله‌خشک} = 4.26 - 0.000099\ W\ EC - 0.388\ W\ pH - 0.00779\ sand - 0.001450\ K$$

هدایت الکتریکی آب: W EC اسیدیته آب: W pH درصد شن: sand پتاسیم خاک: K

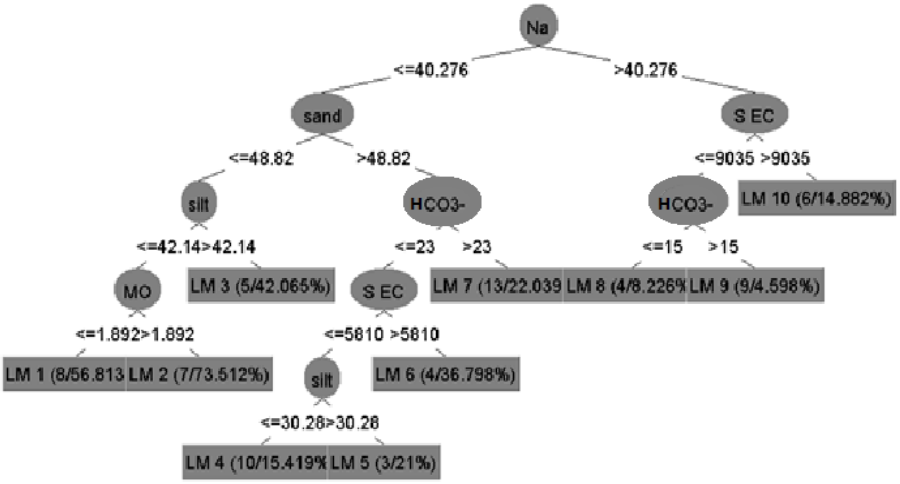


شکل ۲- نمودار پراکنش مقادیر کلاله خشک برآورد شده بر اساس مدل رگرسیونی چند متغیره خطی و مقادیر مشاهداتی

Figure 2 - Distribution diagram of dry stigma values estimated based on multivariate linear regression model and observational values

نمونه‌هایی که سدیم کمتر از ۴۰/۲۷۶ میلی اکی والان گرم بر لیتر بود را در گام بعد با درصد شن بافت خاک به دو دسته‌ی کمتر یا مساوی ۴۸/۸۲ درصد و بیشتر از ۴۸/۸۲ درصد و نمونه‌هایی که سدیم بیشتر از ۴۰/۲۷۶ (mEq/lit) بود را با EC خاک به دوشاخه‌ی کمتر یا مساوی ۹۰۳۵ میکروموس بر سانتی‌متر و بیشتر از آن تفکیک کرد. این تفکیک پارامتری تا جایی ادامه یافت که با ۹ رابطه رگرسیونی برآورد وزن کلاله خشک زعفران حاصل شد (شکل ۳ و جدول ۴). مدل درخت تصمیم M5 در مقایسه با مدل رگرسیون چند متغیره خطی از دقت بالاتری در برآورد وزن کلاله خشک برخوردار بود بطوری که درصد همبستگی را به بیش از ۹۰ درصد (معادل ضریب تبیین ۰/۸۲) بهبود بخشید (شکل ۴). مدل درخت تصمیم M5 نیز به صورت گام به گام اجرا شد و با دخالت دادن پنج متغیر بی‌کربنات آب ($HCo3^-$)، هدایت الکتریکی خاک (S EC)، اسیدیته خاک (pH)، ماده آلی (MO) و درصد سیلت بافت خاک بهترین نتیجه حاصل شد بطوری که مقدار خطای این برآورد RMSE= ۰/۱۲ بدست آمد.

این مدل ابتدا داده‌ها را بر اساس مشاهدات جدول (۴) در مورد موثرترین پارامترهای دخیل در پیش‌بینی وزن کلاله خشک زعفران، حاکی از آن است که پارامترهای شوری خاک (S EC)، یون بی‌کربنات آب (HCo3-)، کلر خاک، درصد ماده آلی خاک (Mo) و درصد سیلت (Silt) بافت خاک، به عنوان موثرترین متغیرها در تمامی ۹ رابطه رگرسیونی مدل درخت تصمیم حضور داشتند (جدول ۴).



شکل ۳- مدل درخت تصمیم برآورد وزن کلاله خشک

Figure 3- Decision tree model for estimating dry stigma weight

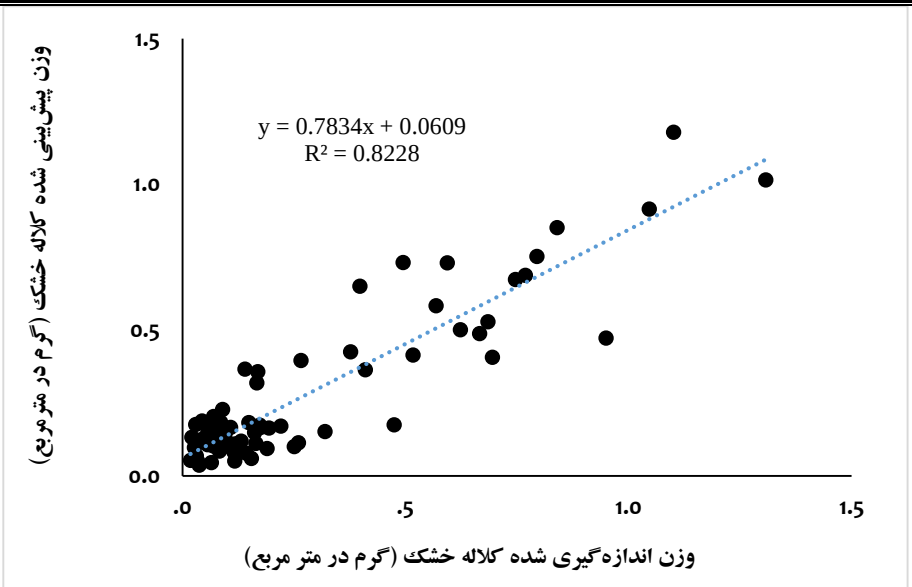
(بی‌کربنات آب ($HCo3^-$)، هدایت الکتریکی خاک (S EC)، ماده آلی (MO) و درصد سیلت (silt) و شن (sand) بافت خاک)

جدول ۴- روابط رگرسیونی مدل درخت برآورد وزن خشک کلاله

Table 4- Tree model regression relationships for stigma dry weight estimation

شماره	رابطه
۱	کلاله خشک = -0.0965 * S pH + 0 * S EC + 0.0101 * HCO3- + 0.0052 * silt - 0.3271 * MO - 0.0049 * Cl + 1.6284
۲	کلاله خشک = -0.0965 * S pH + 0 * S EC + 0.0101 * HCO3- + 0.0048 * silt - 0.3299 * MO - 0.005 * Cl + 1.6177
۳	کلاله خشک = -0.1447 * S pH + 0 * S EC + 0.0129 * HCO3- + 0.0124 * silt - 0.3262 * MO - 0.0047 * Cl + 1.6539
۴	کلاله خشک = 0 * S EC + 0.0039 * HCO3- + 0.0177 * silt - 0.127 * MO - 0.0206 * Na - 0.001 * Cl + 0.5455
۵	کلاله خشک = 0 * S EC + 0.0039 * HCO3- + 0.0182 * silt - 0.127 * MO - 0.0207 * Na - 0.001 * Cl + 0.5432
۶	کلاله خشک = 0 * S EC + 0.0039 * HCO3- + 0.018 * silt - 0.1705 * MO - 0.0221 * Na - 0.001 * Cl + 0.6458
۷	کلاله خشک = 0 * S EC + 0.0039 * HCO3- + 0.014 * silt - 0.127 * MO - 0.0101 * Na - 0.001 * Cl + 0.34
۸	کلاله خشک = 0 * S EC + 0.0018 * HCO3- + 0.0074 * silt - 0.0999 * MO - 0.0019 * Cl + 0.1444
۹	کلاله خشک = 0 * S EC + 0.002 * HCO3- + 0.0074 * silt - 0.0999 * MO - 0.0019 * Cl + 0.1358
۱۰	کلاله خشک = 0 * S EC + 0.0027 * HCO3- + 0.0074 * silt - 0.0999 * MO - 0.0019 * Cl + 0.1274

اسیدیته خاک: S pH، بی کربنات: HCO3، سیلت: silt، ماده آلی: MO ، سدیم: Na ،کلر خاک :CL، هدایت الکتریکی آب: W EC، هدایت الکتریکی خاک: S EC، پتاسیم خاک: K

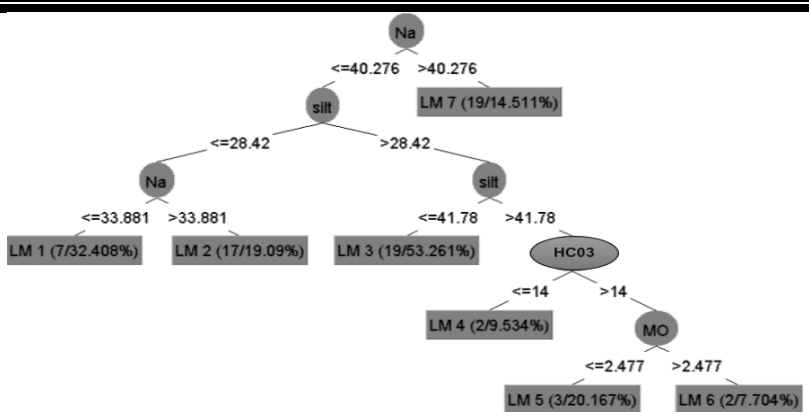


شکل ۴- نمودار پراکنش مقادیر وزن خشک کلاله برآورد شده از مدل درخت تصمیم و مقادیر مشاهداتی

Figure 4 - Distribution diagram of stigma dry weight values estimated from the decision tree model and observational values

تحلیل درخت تصمیم M5 در برآورد وزن گل

در برآورد وزن گل نیز مدل درخت تصمیم M5 با تفکیک داده‌ها به ۶ زیرمجموعه، و ارائه ۶ رابطه رگرسیونی چند متغیره خطی، دقت برآورد را نسبت به مدل رگرسیون چند متغیره خطی بهبود داد بطوری که درصد همبستگی بین مقادیر برآورد شده و اندازه‌گیری شده وزن گل به بیش از ۹۰ درصد (ضریب تبیین معادل ۰/۸۲) رسید و خطای این برآورد تا RMSE=۰/۰۰۹۴ کاهش یافت (شکل ۵ و ۶ و جدول ۵). این نتایج نیز با اجرای گام به گام مدل درخت تصمیم M5 و با دخالت دادن متغیرهای بی‌کربنات، سیلت، ماده آلی و سدیم خاک، نتایج خوبی حاصل شد. مدل درخت تصمیم M5 که بصورت باینری (دودویی) اجراء می‌شود ابتدا براساس مقدار سدیم خاک، مجموعه داده‌ها را به دو گروه دسته‌بندی کردبرای این منظور، نمونه‌هایی که سطح سدیمی آن‌ها بیشتر از ۴۰/۲۷۶ میلی‌اکی‌والان گرم بر لیتر بود، رابطه رگرسیونی شماره (۷) را ارائه کرد. در این رابطه علاوه بر سدیم خاک، یون بی‌کربنات، درصد سیلت بافت خاک و ماده آلی خاک تأثیرگذار بودند. برای سدیم بیشتر از ۴۰/۲۷۶ میلی اکی والان گرم بر لیتر نیز در گام بعدی نمونه‌های باقیمانده براساس درصد سیلت بافت خاک کمتر، مساوی یا بیشتر از ۲۸/۴۲ درصد، به دو دسته دیگر تقسیم شدند و به همین ترتیب تفکیک نمونه‌ها با پارامترهای مقدار سدیم خاک و درصد سیلت بافت خاک، یون بی‌کربنات آب و ماده آلی بافت خاک صورت گرفت تا سرانجام با برقراری (۷) رابطه رگرسیونی، وزن گل در واحد متر مربع برآورد شد (شکل ۵ و جدول ۵). بر اساس جدول (۵)، پارامترهای ماده آلی خاک (Mo)، بی‌کربنات آب (-HCo3)، درصد سیلت بافت خاک (Silt) و سدیم خاک (Na)، مهمترین پارامترهایی هستند که در تمامی روابط رگرسیونی وجود داشتند.



شکل ۵- مدل درخت تصمیم برآورد وزن گل

Figure 5- Tree model of estimating flower weight

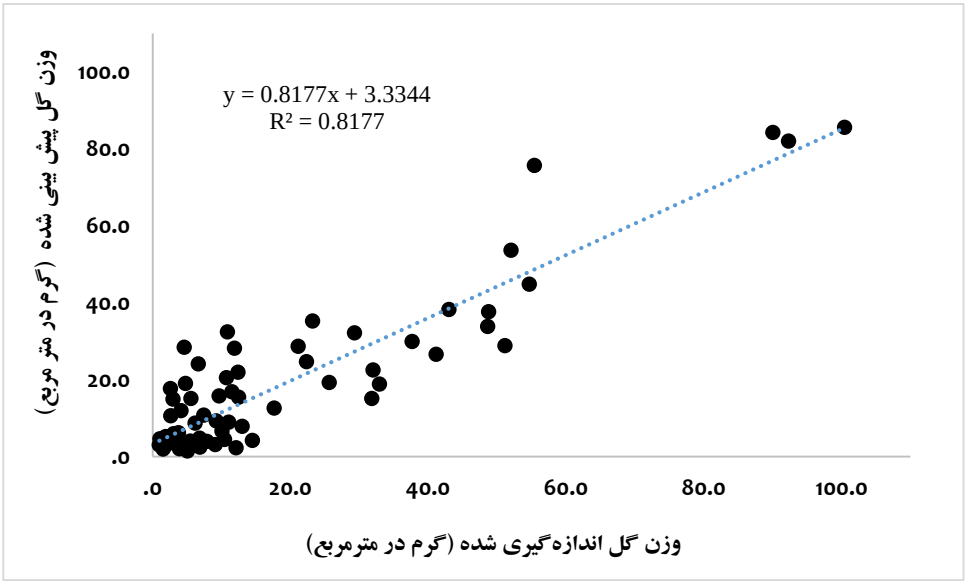
(بی کربنات آب ($HCo3^-$)، ماده آلی (MO)، سدیم خاک (Na)، و درصد سیلت (silt) بافت خاک)

جدول ۵- روابط رگرسیونی مدل درخت تصمیم M5 جهت برآورد وزن گل

Table 5- Regression relation of M5 decision tree model for estimating flower weight

شماره	رابطه
۱	Tot.flo = 0.3465 * HCO3- + 0.8538 * silt - 8.5472 * MO - 0.8214 * Na - 0.1842 * K + 35.2495
۲	Tot.flo = 0.3465 * HCO3- + 0.8538 * silt - 8.5472 * MO - 0.6878 * Na + 21.8159
۳	Tot.flo = -6.0711 * S pH - 0.0286 * HCO3- - 0.529 * Dip.pla + 1.5243 * silt - 15.012 * MO - 0.3773 * Na + 73.9694
۴	Tot.flo = -9.3826 * S pH + 0.9296 * HCO3- + 1.903 * silt - 18.6924 * MO - 0.3773 * Na + 69.1848
۵	Tot.flo = -9.3826 * S pH + 0.885 * HCO3- + 1.903 * silt - 20.3804 * MO - 0.3773 * Na + 76.1491
۶	Tot.flo = -9.3826 * S pH + 0.885 * HCO3- + 1.903 * silt - 20.4797 * MO - 0.3773 * Na + 76.1656
۷	Tot.flo = 0.1675 * HCO3- + 0.7005 * silt - 5.2621 * MO - 0.101 * Na - 0.6775

اسیدیته خاک: S pH، بی کربنات: HCO3، سیلت: silt، ماده آلی: MO ، سدیم: Na ،کلر خاک :CL، هدایت الکتریکی آب: W EC، هدایت الکتریکی خاک: S EC، پتاسیم خاک: K



شکل ۶ - نمودار پراکنش وزن گل برآورد شده از مدل درخت تصمیم M5 و مقادیر مشاهداتی

Figure 6 - Diagram of flower weight distribution estimated from M5 decision tree model and observational values

بحث

ارزیابی مدل‌ها

مدل‌های ساخته شده به هر دو روش رگرسیون گام به گام و درخت تصمیم M5، با استفاده از داده‌های جمع آوری شده‌ی یکسانی بود. که متغیرهای مستقل (نیتروژن، فسفر، پتاسیم، منیزیم، کلسیم، کلر، سدیم، ماده آلی خاک، pH آب و خاک، EC آب و خاک، عمق کاشت، سن مزرعه ودرصدهای سیلت، شن و رس بافت خاک و بی‌کربنات آب)، به عنوان ورودی‌های هر دو مدل رگرسیونی و درخت تصمیم M5، در نظر گرفته شد و متغیرهای وابسته وزن کلاله خشک و وزن گل زعفران به عنوان خروجی‌های هر دو مدل در نظر گرفته شد. معادلات بدست آمده از مدل رگرسیونی با مدل تصمیم درخت M5 ارزیابی شد. مقادیر پیش‌بینی شده با مقدار واقعی جمع آوری شده مقایسه شد و مقادیر معیارهای پراکندگی شامل میانگین خطا، ریشه میانگین خطا و مقدار ضرایب تبیین و همبستگی محاسبه شد.

جدول ۶ - بررسی مقادیر پیش بینی شده در هر دو مدل رگرسیونی و مدل M5

Table 6 - Investigation of predicted values in both regression model and M5 model

ردیف	پارامترها	وزن کلاله خشک زعفران	وزن گل زعفران
۱	مقدار خطای مدل رگرسیونی (gr/m^2)	0.23	16.37
۲	مقدار خطای مدل M5 (gr/m^2)	0.12	9.4
۳	ضریب همبستگی مدل رگرسیونی (٪)	70.0	74.0
۴	ضریب همبستگی مدل M5 (٪)	90.0	90.0
۵	ضریب تبیین مدل رگرسیونی (٪)	49.8	54.8
۶	ضریب تبیین مدل M5 (٪)	82.0	81.0

ارزیابی دو مدل استفاده شده در برآورد وزن کلاله خشک

با برآورد کلاله خشک زعفران در کاربرد مدل رگرسیون چند متغیره خطی، نتایج رضایت بخشی حاصل نشد بطوری که حداکثر ضریب تبیین و ضریب همبستگی بدست آمده در گام چهارم به ترتیب، کمتر از ۴۰ درصد و ۶۲ درصد و حداکثر مقدار خطا برابر RMSE=۰/۲۳ بود. زمانی که تمامی پارامترها وارد مدل شدند، مقدار ضرایب تبیین و همبستگی به ترتیب برابر ۴۹/۸ و ۷۰ درصد شد. براین اساس، مدل رگرسیون گام به گام طراحی شده، توانست پیش‌بینی وزن کلاله خشک زعفران به میزان همبستگی ۷۰ درصد در گام پایانی را داشته باشد.

نتایج حاصل از جدول (۳) به روش رگرسیون، نشان داد که پارامترهای pH و EC آب، درصدشن بافت خاک و پتاسیم، چهار متغیرمهم تاثیرگذار بر وزن کلاله خشک اندازه‌گیری شده بودند که به ترتیب بالاترین ضریب همبستگی وارد مدل شدند (جدول ۳). شوری آب آبیاری موجب افزایش شوری خاک می‌گردد (*Houri et al., 2015*). pH آب آبیاری نیز نقش مهمی درتغییرات pH خاک دارد (*Waraich et al., 2011*). مطالعات زیادی بر روی پتاسیم به عنوان عنصر مهم بعد از نیتروژن که در بسیاری از فعالیت‌های گیاهی نقش دارد انجام گرفته است. کنترل باز و بسته شدن روزنه‌های برگ، تنظیم میزان تنفس و تبخیر در گیاه، از عناصر پرتحرک بوده که در شرایط کمبود از بافت‌های پیر به بافت‌های جوان و مریستمی منتقل می‌شود (*Marschner,1995*). عنصری ضد تنش،که با بالا بردن مقاومت گیاه به انواع تنش، عملکرد گیاه را بهبود می‌دهد. همچنین تنظیم پتانسیل اسمزی سلول، از دیگر وظایف پتاسیم است (*Saeed Rad & Mokhtarian., 2010*). درصد اجزای بافت خاک از عوامل مهم در افزایش عملکرد گل در گیاهان غده‌ای است (*Ikram et al.; 2012*). مطالعاتی که گریستا و همکاران (*Gresta et al,. 2008*)، بهدانی و همکاران (*Behdani et al., 2005*) و *کافی* (*Kaf,i 2002*)، بر *روی با*فت خاک انجام دادند مشخص شد که به دلیل افزایش نفوذپذیری، خاک‌هایی با بافت سبک لومی شنی و لومی رسی سبب افزایش عملکرد زعفران می‌شوند.درصد شن بافت خاک به دلیل ایجاد فضاهای

خالی بزرگتر در ساختمان خاک، نفوذپذیری خاک را بهبود می‌بخشد (*Mahmoodabadi & Mazaheri,2012*). با توجه به تفسیر بیان شده، مدل درخت M5، با دخالت دادن ۵ متغیر بی‌کربنات آب ($HCo3^-$)، هدایت الکتریکی خاک (S EC)، اسیدیته خاک (pH)، ماده آلی (MO) و درصد سیلت بافت خاک بهترین نتیجه حاصل شد. این مدل توانست با ضریب همبستگی ۹۰ درصد و ضریب تبیین بیشتر از ۸۲ درصد، وزن کلاله خشک را براساس پارامترهای فیزیکی شیمیایی آب و خاک برآورد کند (شکل ۴). بطوری‌که در این برآورد حداکثر مقدار خطا برابر با MSE=۰/۱۲ بدست آمد. با توجه به نتایج بیان شده، مدل درخت توانست با کمترین داده بالاترین پیش‌بینی را نسبت به مدل رگرسیونی داشته باشد(جدول ۶). ارزیابی تحقیقی که گیسن و کمپیچلر (۲۰۰۷) در منطقه تاباسکو واقع در جنوب شرقی مکزیک بر روی فرسایش آبی سطحی و زیر سطحی، با مدل تصمیم درخت انجام دادند، حاکی از موفقیت این مدل در پیش‌بینی چگونگی وقوع فرسایش‌های تونلی و چاله‌ای بود. در مطالعه‌ای که (*Taghizadeh Mehrjerdi et al., 2014*) انجام دادند، با استفاده از انواع روش‌های داده‌کاوی پیش‌بینی مکانی گروه‌های بزرگ خاک را در منطقه اردکان یزد به وسعت ۷۲۰۰۰ هکتار، مورد بررسی قراردادند. این نتایج نشان داد که مدل پیش‌بینی شده توسط درخت تصمیم با دقت ۷۰ درصد، عملکرد بهتری نسبت به روش‌های شبکه عصبی با ۶۵ درصد، شبکه عصبی_الگوریتم ژنتیک با ۶۵ درصد، رگرسیون چند جمله‌ای لاجستیک با ۵۵ درصد و آنالیز تشخیصی با ۴۷ درصد داشت.

۴-۳-۴- ارزیابی دو مدل استفاده شده در برآورد وزن گل زعفران

نتایج رگرسیون گام به گام چند متغیره خطی در برآورد وزن گل اندازه گیری شده در زمان گلدهی زعفران برای طراحی یک مدل جامع، قانع کننده نبود. بطوری که در ۴ گام انجام شده، حداکثر ضریب تبیینی که به دست آمد از ۴۵/۱۷ درصد تجاوز نکرد و حداکثر مقدار خطا برابر RMSE=۱۶/۳۸ بود. با این حال با افزودن تمامی پارامترهای اندازه‌گیری شده به مدل، ضرایب تبیین و همبستگی در پایان برآورد وزن گل، به ترتیب برابر ۵۴/۸ درصد و۷۴ درصد شد (جدول ۲). نتایج حاصل از جدول (۲) حاکی از آن است که در بین متغیرهای اندازه‌گیری شده، بی‌کربنات و EC آب، به همراه ماده آلی خاک و درصد سیلت بافت خاک، مهمترین داده‌های ورودی به مدل رگرسیون گام به گام بودند (جدول ۲). از شاخص‌های کیفی خاک به ماده آلی خاک می‌توان اشاره کرد. این شاخص باعث بهبود فرایندهای فیزیکی و شیمیایی و بیولوژیکی خاک است (*Heinz., 1996*). مطالعات انجام شده همبستگی بالایی را بین عملکرد گل و کلاله زعفران با ماده آلی خاک نشان داده است. بهبود تغذیه بنه زعفران و کیفت خاک را بواسطه ماده آلی محتمل دانست. از منابع مهم تاثیر گذار بر تغذیه زعفران کودهای دامی، کاه و کلش، کمپوست و... (FYM[6])می‌باشد (*Koocheki et al., 2006*). مدل درخت تصمیم M5، پس از انجام عملیات مدل‌سازی و ایجاد روابط رگرسیونی بین پارامترهای در نظر گرفته شده، توانست با دقت بالایی مدلی را پیش‌بینی کند که در آن ضریب همبستگی بیشتر از ۹۰ درصد و ضریب تبیین معادل ۸۱ درصد شد (جدول ۶، شکل ۶). به‌طوری‌که در این برآورد حداکثر مقدار خطا برابر با RMSE=۰/۰۰۹۴ بدست آمد. در این مدل‌سازی که با روش گام به گام صورت گرفت با دخالت دادن پارامترهای تاثیرگذار درصد سیلت خاک، ماده آلی و سدیم خاک، نتایج خوبی حاصل شد. نمک یا شوری که حاصل ترکیب دو یون کلر و سدیم است آثار تخریبی بر رشد و عملکرد گیاه دارد. در واقع سمیت این یون‌ها را دلیل این فرایند بیان کرده‌اند.آثار مخرب این یون‌ها و البته برخی یون‌های دیگر که مستقیم یا غیر مستقیم باعث دفرمه شدن ساختمان مولکول‌های بزرگ پروتئین و به هم ریختگی غشاء سیتو پلاسمی و سلولی شده (نشت یونی) به اثبات رسیده‌است (*Saadati et al., 2004*).

با بررسی پارامترهای تاثیر گذار در هر دو مدل مشاهده شد که از چهار پارامتری که در گام چهارم مدل رگرسیونی و مدل درخت تصمیم M5، مورد استفاده قرار گرفت، سه پارامتر، ماده آلی خاک، سدیم خاک و درصد سیلت بافت خاک، مشترک بوده و در هردو مدل استفاده شد. مشاهدات نتایج هر دو مدل گویای برتری مدل درخت M5 می‌باشد(جدول ۶).

[6] Farm yard manure

مطالعات انجام شده در ارزیابی باروری ابرها و پیش‌بینی حجم بارش در استان فارس بیانگر دقت بالای مدل درخت تصمیم بود (*Shaykh, 2011*).

نتیجه‌گیری

گرانی نهاده‌های کشاورزی، صرف زمان و هزینه‌های مختلف در مراحل مختلف کاشت، داشت و برداشت از یک سو و کاهش هزینه‌های تولید و کسب عملکرد بهینه، همواره مورد توجه کشاورزان بوده و هست. پیشرفت‌های روزافزون علمی این امکان را می‌توانند در اختیار صنعت کشاورزی قرار دهد تا با استفاده از جمع‌آوری داده‌های زودیافت آب و خاک و امکانات سخت‌افزاری و نرم‌افزاری، در جهت کاهش هزینه‌های تولید گام برداشته و کشاورزان را از آینده عملکرد محصول مطلع سازید. در تحقیق حاضر ۱۲ نمونه آب آبیاری که از چاه‌ها و قنوات منطقه مورد مطالعه و ۶۹ نمونه خاک از ۶۹ مزرعه منطقه در سال ۹۷-۹۸ جمع-آوری شد و با استفاده از پارامترهای فیزیک و شیمیایی آب و خاک جمع‌آوری شده از این نمونه‌ها، انتخاب مدلی برتر مورد مطالعه قرار گرفت. در این پژوهش، مقایسه دو مدل رگرسیون گام به گام و مدل درخت تصمیم M5 صورت گرفت.

با توجه به نتایج، مدل درخت تصمیم M5 از توانمندی بالایی در جهت پیش‌بینی وزن گل و وزن کلاله خشک برخوردار بود. به‌طوری که وزن کلاله خشک و وزن گل را با ۹۰ درصد همبستگی در انتهای مدل‌سازی و برای منطقه مورد مطالعه با توجه به داده‌های وارد شده، برآورد کرد. در حالی که مدل برآوردی رگرسیونی گام به گام در ایده‌آل‌ترین حالت و ورود تمامی پارامترهای(۱۸) اندازه گرفته شده توانست وزن کلاله خشک و وزن گل را به ترتیب با همبستگی‌های ۷۰، ۷۴ درصد، پیش‌بینی کند.

تاثیرگذارترین پارامترها در پیش‌بینی‌های انجام شده با درخت تصمیم M5، برای مدل‌سازی وزن گل و کلاله خشک زعفران شامل، EC آب و خاک، بی‌کربنات آب، سدیم و پتاسیم خاک، ماده آلی خاک و درصد شن و سیلت بافت خاک بود. که اکثر این پارامترها را می‌توان با چند آزمایش ساده اندازه‌گیری کرد. به نظر می رسد که در تحقیقات مشابه این پژوهش، ساده و قابل فهم بودن مدل درخت تصمیم M5 را، جایگزین مناسبی برای سایر روش‌های داده‌کاوی می‌دانند. این مدل، با استفاده از روابط ساده‌ای که در محدوده‌ی خاصی از داده‌های ورودی می‌دهد، درک بهتری از خروجی را خواهد داشت.

منابع

Abrishami, M.H. 2004. Saffron from long ago to today. Amirkabir Publications. Tehran p.119.

Behdani, M. A., Koocheki, A., Nassiri Mahallati, M., and Rezvani Moghaddam, P. 2005. Evaluation of quantitative relationships between saffron yield and nutrition (on farm trial). Journal Journal of Field Crops Research 3(1):1-14. (In Persian with English Summary).

Beheshti, A., and Faravani, M. 2003. Investigation the effect of different planting proportions and densities on yield and yield components of saffron and Caraway mixed cropping. 3th National Congress of Saffron, Mashhad, Iran, 2-3 December.

Bremner, J. S., & Mulvaney, C. S. (1982). Nitrogen-total. In A. L. Page (Ed.), Methods of Soil Analysis, Part American Society of Agronomy (pp. 595-624). Madison, Wisconsin.

Cichosz P. Data Mining Algorithms: Explained Using R. New York: John Wiley & Sons; 2014.

Dosti Irani, A. And Golzarian, M.R. 2016. A review of the applications of artificial intelligence in agricultural processes. 10th National Congress of Biosystem Mechanics (Agricultural Machinery) and Mechanization of Iran.

Falahi, M., Varvani, H. And Golian, s. 2011. Prediction of rainfall using tree regression model to control floods. Fifth National Conference on Watershed Management and Management of Water and Soil Resources, Kerman.

Ghorbani Khalil, Sohrabian Elahe, Jazi Meysam. 2016. Evaluation of hydrological methods and data mining in simulation and forecast of monthly flow rate (Case study: Arazkoseh hydrometric station). Water and Soil Conservation

Research (Agricultural Sciences and Natural Resources), Volume 23, Number 1, Pages 203-217.

Geissen V., Kampichler C. 2007. Superficial and subterranean soil erosion in Tabasco, tropical mexico.

Gresta, F., Lombardo, G.M., Siracusa, L., and Ruberto, G. 2008. Saffron, an alternative crop for sustainable agricultural systems, a review. Agronomy Sustainable Development 28: 95-112.

Haynes, R.J. 1996. Labile organic matter fraction as central components of the quality of agricultural soils. Advances in Agronomy. 85: 221-261.

Hosseini, R., Sarmad, M., Jabbari Noghabi, M. 2013. Data Mining in R Using the Ratlle Package. Statistical Thought of the Eighteenth Year, No. Olf No. 35, pp. 29-17.

Hoori,M.A., Naseri, A., Boroumandansab, S. A. and Kiani, A. 2015. The effect of under-irrigation and irrigation water salinity on soil salinity distribution and vegetative growth of date seedlings. Journal of Soil and Water Resources Protection. 13-1: 14

Ikram, S., Habib, U., and Khalid, N. 2012. Effect of different potting media combination growth and vase life of tuberose (Polianthes tuberose Linn.). Pakistan Journal of Agricultural Sciences 49(2): 121-125.

Kafi, M. 2002. Saffron, Production and Processing. Ferdowsi University of Mashhad Press. 276p. (In Persian).

Kalantari, Kh. 2008. Data Processing and Analysis in Socio-Economic Research, Third Edition, Farhang Saba Publications.

Koocheki, A., Nassiri, M., Behdani, M.A., 2006. Agronomic attributes of saffron yield at fertilizers. J. Med. Plant. 40(4), 58-68.

Luo, Z., Yaolin, L., Jian, W., and Jing, W. 2008. Quantitative mapping of soil organic material using field spectrometer and hyperspectral remote sensing. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 37: 901-906.

Mahmoodabadi, M., and Mazaheri, M.R. 2012. Effect of some soil physical and chemical properties on permeability in field conditions. The Iranian Society of Irrigation and Water Engineering 2(8): 14-25. (In Persian with English Summary).

Mahmoudabadi, E., and Karimi, A. 2015. Mapping of calcium carbonate equivalent and clay content of surface soil using geostatistical methods (Case study: Chitgar park, Tehran). RS GIS Tech. Nat. Resour. 6(3): 73-85. (In Persian).

Maleki,F., Kazemi, H., Siahmarguee, A., and Kamkar, B., 2019. Evaluation of Saffron Cultivation Development in AzadshahrTownship (Golestan Province) with Emphasis on Climatic Parametrs, Journal of Saffron Research. 7(1), 123-143. (In Persian).

Marschner, H. 1995. Mineral nutrition of higher plants. Second edition, Academic press, London.

Montgomery, D.C., Peck, E.A., and Vining, G.G. 2015. Introduction to Linear Regression Analysis. John Wiley & Sons

Movahedi Naini, A. And Rezaei, M. 2008. Soil Physics (Basics and Application). Publications of Gorgan University of Agricultural Sciences and Natural Resources, p. 474.

Munshi, A.M. 1994. Effect of N and K on the floral yield and corm production in saffron under rainfed condition. Indian Arecant and Spices Journal. 18: 24-44.

Nelson, R. E. 1982. Carbonate and gypsum. In: Methods of Soil Analysis. Part II. Page, A. L. (Ed) American Society of Agronomy, Madison, Wisconsin, USA.

Olsen, S.R., Cole, C.V., Watanabe, F.S., and Dean, L.A. 1954. Estimation of available phosphorus in soils by extraction with sodium bicarbonate. USDA Circular 939, US Gov. Printing Office, Washington, DC.

Page, A., Miller, R., and Keeney, D. 1982. Methods of Soil Analysis.2th ed. Part2: Chemical and biological properties. Soil Sci. Soc. Am. Inc. publisher.

Page. A.L., Miller, R.H. and Jeeney, D.R. 1992. Methods of Soil Analysis, Part 2. Chemical and mineralogical properties. SSSA Pub., Madison. 1159p.

Quinlan, J .R. 1992 .Learning with continuous classes. In proceedings AI, 92 (Adams & Sterling, Eds), 343-348, Singapore: World Scientific.

Remesan R., Shamim M.A., and Han D. 2008. Model data selection using gamma test for daily solar radiation estimation. Hydrological Processes 22: 4301-4309.

Rezvani Moghaddam, P., Khorramdel, S., and Mollafilabi, A., 2015. Evaluation of soil physical and chemical

characteristics impacts on morphological criteria and yield of saffron (Crocus sativus L.). Journal of Saffron Research. 3(2): 188-203. (In Persian).

Richards L. A. 1954. Diagnosis and Improvement of Saline and Sodic Soils. U.S.D.A Agric. Handb. NO. 60 Washingtin DC.

Rivandi, M., Ghasemnezhad, A., Hemmati, K.h.,Ghorbani, k.h., and Abhari, A., 2020. Effects of pH, Bicarbonate and Salinity of Irrigation Water on Yield and Flowers Yield Components of Saffron. Journal of Saffron Research, 8(2), 241-253. (In Persian).

Saadati, A., Ghare Yazi, B., and Mirmohammadi Meybodi, A.M. 2004. Salt stress and physiological aspects of plant breeding. University of Esfahan. (In Persian).

Saeedi Rad M. and Mokhtarian A. 2010. Scientifitic principles of cultivation, growing and harvesting of saffron.

Salehi Sadeghian, J. And. Ebrahimi, A. 2002. Advanced Statistical Analysis; Hava Tazeh publications.

Sattari, M.T., Pal, M. Apaydin, H. Ozturk, F, 2013. M5 Model Tree Application in Daily River Flow Forecasting in Sohu Stream, Turkey, Water Resources. 40(3): 233–242.

Shaykh, M.A. 2011. Evaluation of Cloud Fertility Projects Using Regression Decision Tree Model (Case Study: Central Iran - Fars Province), Master Thesis in Civil Engineering, University of Water and Power Industry, Tehran.

Taghizadeh Mehrjerdi, R. Sarmadian, F. Omid, M. Tumanian, N. Rusta, M.J. Rahimian, M.H. 2014. Digital mapping of soil classes using various data mining methods in Ardakan region of Yazd, Journal of Agriculture, Agricultural Engineering, Volume 37, Number 2.

Terzi Ö. 2007. Data mining approach for estimation evaporation from free water surface. Applied Sciences. 7(4): 593-596.

Waraich, E. A., Saifullah, R. A., Ehsanullah, M. Y. 2011. Role of mineral nutrition in alleviation of drought stress in plants. Australian Journal of Crop Science. 5(6): 764-777.

Journal of Saffron Research (semi-annual)

Vol.9, No.2, Fall & Winter, 2021-2022

p. 352-367

http://dx.doi.org/10.22077/JSR.2021.4399.1162

University of Birjand

Original Article:

Prediction of Saffron Flower and Stigma Yield Based on the Physical and Chemical Properties of Water and Soil Using Linear Multivariate Regression Models and M5 Decision Tree

Morteza Rivandi [1], Azim Ghasemnezhad [2 *], Khalil Ghorbani [3], Khodayar Hemmati [4], Abbas Abhari [5]

1- M.Sc. student of Gorgan University of Agricultural Sciences and Natural Resources, Iran.

2, 3- Associate Professor of Horticulture, Gorgan University of Agricultural Sciences and Natural Resources, Iran.

4- Associate Professor of Water Engineering, Gorgan University of Agricultural Sciences and Natural Resources, Iran.

5- Assistant Professor, Department of Agriculture, Payame Noor University, 19395-4697 Tehran, Iran.

** Corresponding Author Email: ghasemnezhad@gau.ac.ir*

Received 2 June 2021; Accepted 30 August 2021

Abstract

The exciting growth of various sciences and technologies and the complexity of decision-making in recent decades, have marked new ways for humanity to use information systems and artificial intelligence, accurately and quickly and provide a solution for its time-consuming scientific and technical predictions and calculations. This study was conducted to evaluate the predictive outcome of two stepwise regression models and the M5 decision tree model under the influence of different water and soil properties on saffron flower and stigma yield in 2019 in Sabzevar saffron fields (located at longitude "57.43" Latitude "36.12") and was performed in the laboratory of the Faculty of Plant Production of Gorgan University of Agriculture Sciences and Natural Resources. In April, after the end of the saffron growing season, soil samples from a depth of zero to 30 cm and 12 irrigation water samples from 69 saffron farms were prepared and transferred to the laboratory to analysis and measuring the physical and chemical properties of soil samples. Around 13 parameters including pH, acidity, and percentage of soil components, soil elements, etc. and acidity, bicarbonate, etc. were measured in soil and water samples, respectively. Flowers were collected at the time of flower emergence from the specified areas of the fields and the desired measurements were made. The results showed that, as the soil and water analysis is relatively expensive, the M5 decision tree model has more accuracy due to the speed and lower cost than the regression model. So that, in the result of predicting the stepwise regression model, in the most ideal case and entering all the measured parameters, dry stigma weight and flower weight with correlations of 70 and 74%, respectively, and the error value is 0.23 RMSE and RMSE 16.38 were predicted. While the M5 decision tree model with lower parameters had a high capability to predict flower weight and dry stigma weight. It estimated the weight of dry stigma and flower weight with 90% correlation and error value equal to RMSE = 0.12 and RMSE = 9.4 at the end of modeling for the study area. Therefore, the M5 decision tree method is recommended in evaluating and predicting various factors on saffron yield.

Keywords: Artificial intelligence, Modeling, Saffron yield, Soil elements.