

تخمین محصول سیب زمینی با استفاده از تصاویر ماهواره‌ای سنتینل-۲، مطالعه موردی: شهرستان سراب

محمد رسولی نیا^۱، علیرضا شریفی^{۲*}

^۱ دانشجوی کارشناسی ارشد مهندسی نقشه‌برداری، سنجش‌ازدور - دانشکده مهندسی عمران - دانشگاه تربیت دبیر

شهید رجایی

mohamadrasolinia@gmail.com

^۲ استادیار گروه مهندسی نقشه‌برداری - دانشکده مهندسی عمران - دانشگاه تربیت دبیر شهید رجایی

a_sharifi@sru.ac.ir

(تاریخ دریافت اسفند ۱۳۹۹، تاریخ تصویب خرداد ۱۴۰۰)

چکیده

کشت سیب‌زمینی به‌عنوان یکی از محصولات کشاورزی جایگاه ویژه‌ای در سبد غذایی انسان‌ها دارد و پس از ذرت و گندم، سومین منبع غذایی در جهان است. مدل‌های سنتی کشت سیب‌زمینی محدودیت‌هایی از قبیل هزینه زیاد، به دست آوردن داده ورودی مورد نیاز و کمبود اطلاعات مکانی در برخی موارد باعث شد که این مدل‌ها به فراموشی سپرده شوند. با تکیه بر علوم نوین و فناوری‌های جدید، پیشرفت‌های چشمگیری در زمینه تولید و کیفیت محصولات زراعی رخ داده است. هدف از انجام این پژوهش مشخص نمودن بهترین بازه زمانی باهدف تولید مدل‌های تخمین برای محصول سیب‌زمینی و همچنین تعیین بهترین الگوریتم برای مدل‌سازی و در نهایت دستیابی به بهترین دقت برای پیش‌بینی میزان محصول تولیدی و کیفیت آن هست. در کنار این موارد تأثیر عوامل زمینی مانند آبیاری، سم‌پاشی، کود دهی و میزان بارش نیز به‌عنوان فاکتورهای مؤثر مورد بررسی قرار گرفته است. هدف از این تحقیق ادغام داده‌های ماهواره‌ای و داده‌های زمینی بر روی ۴۰ سایت کشت سیب‌زمینی در شهرستان سراب استان آذربایجان شرقی جهت تخمین محصول قابل برداشت است. از تصاویر ماهواره‌ای سنتینل-۲ که برای استخراج شاخص NDVI اراضی مورد مطالعه استخراج شد و میزان همبستگی آن‌ها با داده‌های زمینی به دست آمد. بیشترین همبستگی در تاریخ ۲۸ خرداد به میزان ۰,۴۸ بود. دلیل پایین بودن این همبستگی غده بودن سیب‌زمینی و تأثیر کم سبزی‌نگی روی محصول تولیدی می‌باشد. برای مدل‌سازی از چهار الگوریتم یادگیری ماشین استفاده شد که روش فرآیند گوسی بهترین نتایج را به‌همراه داشت ($R^2=0.61$). نتایج نشان داد که برای افزایش دقت مدل‌سازی و تولید مدل‌های تخمین می‌توان از داده‌های ماهواره‌ای با رزولوشن بهتر استفاده کرد تا نتایج بهتری تولید شود و همچنین می‌توان تعداد اراضی که به‌عنوان داده آموزشی برای مدل‌سازی استفاده می‌شدند را بالا برد تا مدلی دقیق تولید گردد.

واژگان کلیدی: تخمین محصول سیب‌زمینی، سنجش از دور، سنتینل-۲، شاخص طیفی، رگرسیون.

* نویسنده رابط

۱- مقدمه

در سال‌های اخیر جمعیت جهان رو به افزایش بوده که در سال ۲۰۱۸ به حدود ۷/۷ میلیارد نفر رسیده است. در پنج سال آینده جمعیت جهان حدود ده میلیارد نفر خواهد شد و بدین ترتیب تقاضا برای غذا، منابع آب و زمین‌های حاصلخیز برای تولید مواد غذایی افزایش خواهد یافت (۱). از این‌رو کشاورزی دقیق در تلاش است تا با استفاده از فناوری‌های جدید زمینه بهبود کمیت و کیفیت محصولات کشاورزی را ایجاد کند (۲). کشاورزی دقیق یک مفهوم جدید در کشاورزی امروزی است و بر مبنای وجود ناهمگونی در سطح مزرعه استوار است. یکی از دلایل رشد کشاورزی دقیق در میان دانشمندان و کشاورزان، پیشرفت تکنولوژی در زمینه‌های مختلف از جمله سیستم تعیین مختصات جهانی، سنجنده‌های تصویربرداری و ابزارهای مدیریت اطلاعات مکانی است (۳).

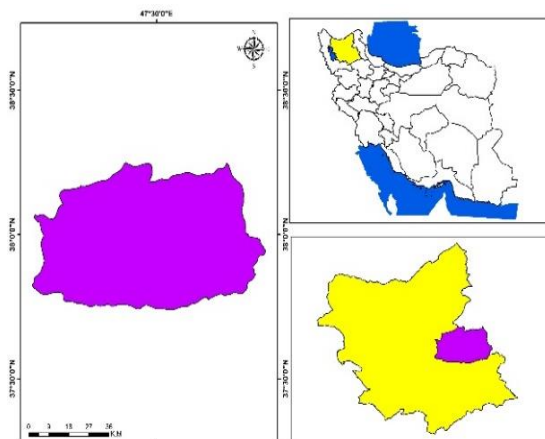
بر این اساس اطلاعات جمع‌آوری شده توسط فناوری‌های ذکر شده برای انجام محاسبات دقیق در مورد سطح ناهمگونی در مزرعه از جهات مختلف از جمله مقدار مواد مغذی خاک، گسترش و پخش آفت‌ها، بیماری‌ها و علف‌های هرز در سطح مزرعه همچنین تصمیم‌گیری در مورد موعد فعالیت‌های مدیریتی و پیش‌بینی مقدار عملکرد، مورد ارزیابی قرار خواهند گرفت (۴). استفاده از سنجنش‌ازدور این امکان را به ما می‌دهد که در مدت‌زمان بسیار کم فضای بسیار بزرگی را کنترل کنیم که این کار هزینه‌های سنگین را کاهش خواهد داد (۵).

در کنار برنج و گندم، سیب‌زمینی سومین محصول غذایی مهم در جهان است که بیش از یک میلیارد نفر از آن تغذیه می‌کنند که هم‌زمان با افزایش جمعیت، تقاضا برای این محصول افزایش یافته است؛ بنابراین نیاز است شیوه‌های بهتر حفاظت و مدیریت محصول به‌منظور بهبود تولید به کار گرفته شود (۶). در سال ۲۰۰۳ سیلان و همکاران معماری دقیق کشاورزی را بر اساس طراحی سامانه‌های فیزیکی-سایبری پیشنهاد دادند که این فناوری شامل سه لایه (لایه فیزیکی، لایه شبکه و لایه تصمیم‌گیری) بود. این معماری می‌تواند برای کاربردهای مختلف کشاورزی استفاده شود (۷). در سال ۲۰۰۵ پاتاک و همکاران موضوع کنترل هوشمند آفت را بر اساس سامانه‌های

فیزیکی-سایبری مطرح کرد که پیشنهادی برای نظارت بر موش‌های زمین‌های کشاورزی بود (۸). در سال ۲۰۰۹، تیملین و همکاران و همکارانش موضوع کنترل کیفیت تحلیلی خاک با تکنیک‌های سنتی و طیف‌سنجی را در برزیل مورد بررسی قرار دادند که اهمیت داده‌های ماهواره-ای در کنار داده‌های زمینی را نشان می‌دهد (۹).

در سال‌های اخیر تحقیقات زیادی در زمینه محصول سیب‌زمینی صورت گرفته است. در سال ۲۰۱۴ مطالعه‌ای توسط بالا و اسلام روی محصول سیب‌زمینی با استفاده از تصاویر لندست ۸ و سنتینل ۲ صورت گرفت. در این تحقیق، برای پیش‌بینی عملکرد محصول سیب‌زمینی از شاخص‌های پوشش گیاهی در طی مراحل رشد سیب‌زمینی در کنار تجزیه و تحلیل فعالیت‌های زیستی استفاده شد. نتایج نشان داد که محصول برداشت‌شده به روش جدید با افزایش ۱۰ تنی در هر هکتار مواجه شده است (۱۰). دیواکس و همکاران در سال ۲۰۱۴ مطالعه‌ای روی دو محصول زیتون و سیب‌زمینی با رویکرد کشاورزی دقیق به ترتیب در یونان و هلند انجام دادند. نتایج نشان داد که بازده محصول در طی یک دوره مشابه نسبت به روش‌های سنتی ۲۱ الی ۲۶ درصد افزایش یافته است (۱۱). در مطالعه دیگر، در سال ۲۰۱۵ منطقه‌ای به وسعت ۵۰ هکتار را مورد ارزیابی قرار دادند و از تصاویر ماهواره‌ای برای استخراج شاخص‌های آب استفاده شد. منطقه از نظر آبیاری به سه سطح تقسیم شد و اهمیت آبیاری صحیح و بهنگام مشخص گردید (۱۲). ژائو و همکاران در سال ۲۰۱۶ نیز مطالعه‌ای روی محصول سیب‌زمینی در کشور اسپانیا صورت گرفت و ۳۳ زمین بزرگ زیر کشت سیب‌زمینی مورد بررسی قرار گرفتند. این منطقه دارای تابستان‌های خشک و زمستان‌های بارانی است و میانگین بارش ۴۳۰ میلی متر در این منطقه گزارش شده است (۱۳).

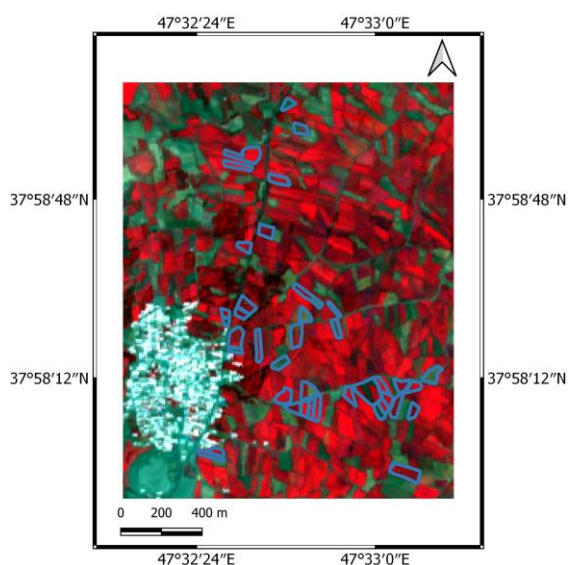
مطالعات اخیر مجموعه‌ای از پژوهش‌هایی می‌باشد که اقدام به بررسی محصولات به‌صورت تک‌بعدی کرده‌اند، یا مطالعات با تکیه بر اطلاعات ماهواره‌ای یا صرفاً با اطلاعات زمینی انجام شده است. در این مقاله ابتدا روش سنتی کشت سیب‌زمینی و روش‌های نوین کشت این محصول مورد بررسی قرار گرفت. سپس با تلفیق داده‌های زمینی و ماهواره‌ای سعی شده است مدلی ارائه گردد که خلأهای مطالعات اخیر پوشش داده‌شده و همچنین سلامت گیاه در طول دوره رشد بررسی و میزان محصول برداشتی پیش‌بینی گردد.



شکل ۱- موقعیت منطقه مورد مطالعه

۲-۲- داده زمینی

در این تحقیق، طی دو سال گذشته ۴۰ زمین کشاورزی در منطقه مورد ارزیابی قرار گرفت. برای این که داده های زمینی دقت مطلوب داشته باشند از روش جمع آوری مستقیم داده استفاده شد. جدول ۱ خلاصه ای از آماره های داده های زمینی را نشان می دهد.



شکل ۲- تصویر سنتینل-۲ از منطقه مورد مطالعه با ترکیب رنگی کاذب (کانال قرمز: NIR، کانال سبز: Red، کانال آبی: Green).

موقعیت اراضی که اطلاعات آن ها برداشت شده با چندضلعی آبی رنگ نشان داده شده است. داده های زمینی با مراجعه به سر زمین های کشاورزی و تکمیل پرسشنامه هایی از کشاورزان و همچنین با مصاحبه هایی که از متخصصین کشاورزی منطقه انجام گرفت، جمع آوری شدند. این اطلاعات شامل نحوه کشت و کشت و

هدف از انجام این پژوهش مشخص نمودن بهترین بازه زمانی باهدف تولید مدل های تخمین برای محصول سیب-زمینی و همچنین تعیین بهترین الگوریتم برای مدل سازی و در نهایت دستیابی به بهترین دقت برای پیش بینی میزان محصول تولیدی و کیفیت آن هست. در کنار این موارد تأثیر عوامل زمینی مانند آبیاری، سم پاشی، کود دهی و میزان بارش نیز به عنوان فاکتورهای مؤثر مورد بررسی قرار می گیرند تا مدلی با بالاترین دقت و کیفیت تولید گردد.

۲- مواد و روش ها

۲-۱- منطقه مورد مطالعه

در این تحقیق اراضی کشاورزی شهرستان سراب (طول جغرافیایی ۴۷ درجه و ۲۳ دقیقه، عرض جغرافیایی ۳۷ درجه و ۵۶ دقیقه) مورد مطالعه قرار گرفته است (شکل ۱). شهرستان سراب به دلیل قرار گرفتن در بین دو رشته کوه بزقوش در جنوب و ارتفاعات سبلان در شمال دارای تابستان معتدل و زمستان سرد است. شواهدی از توسعه روستایی، کشاورزی و دامداری در منطقه سراب در حدود ۸۰۰۰ سال قبل از میلاد وجود دارد (۱۴). این منطقه از مناطق پر آب استان آذربایجان شرقی محسوب می شود و علاوه بر آب های سطحی، ۱۳۵ چشمه و قنات در منطقه وجود دارد. به دلیل خشک سالی اخیر، تعدادی از این چشمه ها و قنات ها خشک شده اند. بیش از ۱۸۰۰ حلقه چاه عمیق و نیمه عمیق در این منطقه وجود دارد و در فصل برداشت در مجموع ۱۶۸ میلیون مترمکعب آب از منابع زیرزمینی برای کشاورزی استفاده می شود (۱۵).

اغلب محصولات زراعی این منطقه شامل گندم، سیب زمینی، سبزیجات، صیفی جات و حبوبات می باشد. همچنین محصولات باغی مانند سیب، گلابی، گردو و زردآلو از جمله تولیدات شهرستان سراب محسوب می شوند. سیب زمینی اصلی ترین و مهم ترین محصول سراب است به نحوی که این شهرستان سالانه ۱۵۰ هزار تن سیب زمینی تولید می کند و بدین ترتیب ۵۰ درصد سیب زمینی استان آذربایجان شرقی و ۴ درصد سیب زمینی کشور در سراب کشت می شود. این محصول به بازارهای عراق و جمهوری آذربایجان صادر می شود و نقشی اساسی در درآمد مردم این منطقه دارد (۱۶).

برداشت محصول، مساحت زمین‌ها، ساعت آبیاری و نحوه آبیاری زمین‌ها، میزان محصول برداشتی در دو سال گذشته، کودها و سموم مورد استفاده، آفت‌ها و روش‌های مقابله با آن‌ها می‌باشد.

۲-۳- داده ماهواره‌ای

در این مقاله از تصاویر ماهواره سنتینل ۲ برای نظارت بر محصول سیب زمینی استفاده شده است (شکل ۲). از ویژگی‌های اصلی این ماهواره می‌توان به تصویربرداری چندمنظوره با ۱۳ باند در طیف مرئی، مادون قرمز نزدیک و موج مادون قرمز کوتاه از طیف الکترومغناطیسی، توان تفکیک مکانی ۱۰، ۲۰ و ۶۰ متری و میدان دید ۲۹۰ کیلومتری اشاره کرد (۱۷).

جدول ۱- پارامترهای آماری منطقه مورد مطالعه.

انحراف معیار	میانگین	بیشینه	کمینه	
	۴۳۹۶	۱۳۵۶۲	۱۳۴۰	مساحت زمین (مترمربع)
	۷,۲۵	۶۱,۸۶	۳۱,۹۷	میزان بهره‌وری (تن بر هکتار)
	۰,۰۳	۰,۸۳	۰,۶۸	NDVI (۲۸ خرداد)

۲-۴- روش پیشنهادی

برای ساخت مدل کل تصاویر منطقه طی ماه‌های اسفند تا مهر در مدت دو سال در نظر گرفته و پیش‌پردازش‌هایی از قبیل حذف ابر و سایر خطاها انجام شد تا تصاویری با حداقل خطا داشته باشیم. سپس شاخص NDVI محاسبه گردید که برای محاسبه آن از باندهای قرمز و مادون قرمز به کار گرفته شد (شکل ۳). با محاسبه شاخص NDVI اطلاعات مفیدی در مورد گیاه از قبیل سلامتی، سرزنده بودن گیاه، وجود یا عدم وجود آفت، آبیاری صحیح به دست آمد. از طرفی اطلاعات زمینی، زمین‌های کشاورزی از طریق پرسشنامه‌هایی گردآوری شده است. برای مدل‌سازی با دو مجموعه داده‌ی زمینی و ماهواره‌ای سروکار داریم. استفاده جداگانه هر یک از آن‌ها با کمبودها و نقص‌هایی همراه خواهد بود؛ بنابراین برای مدل‌سازی از تلفیق اطلاعات زمینی و ماهواره‌ای استفاده گردید که مدلی با حداقل کمبود و نقص ساخته شود. برای مدل‌سازی معمول است که به صورت تکراری کار شود تا

بهترین الگوریتم به دست آید، از این رو رگرسیون خطی برای مدل‌سازی بهترین انتخاب خواهد بود.

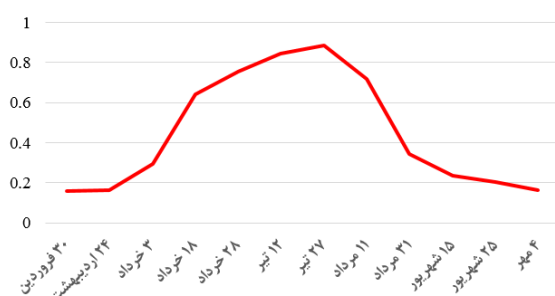
اطلاعات مربوط به سلامت و رشد گیاهان از جمله اطلاعات کاربردی برای کشاورزان می‌باشد چراکه کمبود-های مطرح در رشد گیاه به حداقل می‌رسد و نقاط درگیر تنش در گیاهان شناسایی می‌گردد. این موضوع باعث ایجاد زمینه‌ای برای کشت گیاهانی سالم با حداقل آفت و حداکثر بازدهی می‌شود. در این پژوهش یک ماه قبل از برداشت محصول عملکرد محصول پیش‌بینی شده است. بهترین مدل، مدلی است که عملکرد مناسب‌تری داشته باشد و پیش‌بینی بازده محصول یک ماه قبل از برداشت، نتایج واقع‌بینانه و نزدیک به واقعیت در برداشته است.

رویکردهای یادگیری ماشین با استفاده از یک مدل قابل‌نمایش مستقیم از داده‌ها، رابطه بین ورودی و خروجی را یاد می‌گیرند. ابر پارامترهای مدل به‌طور معمول تنظیم می‌شوند تا خطای پیش‌بینی را در یک مجموعه داده اعتبار سنجی به حداقل برسانند. به این ترتیب، فرد به دنبال بهترین قابلیت‌های تولید است، نه تنها عملکرد خوب در مجموعه آموزش که می‌تواند یک راه حل کاربردی را ایجاد کند. چهار الگوریتم پیشرفته رگرسیون یادگیری ماشین باهدف مقایسه در نظر گرفته شده که عبارت‌اند از رگرسیون خطی، شبکه‌های عصبی (NN)، رگرسیون بردار پشتیبان (SVR) و رگرسیون فرایند گوسی (GPR). همه این روش‌های رگرسیون به لطف آموزش نسبتاً سریع، عملکرد خوب و استحکام نسبت به مشکل بیش‌ازحد، در حوزه‌های مختلف کاربردی محبوب هستند.

الف) رگرسیون خطی

اگر برای شناسایی و پیش‌بینی متغیر وابسته فقط از یک متغیر مستقل استفاده شود، مدل را رگرسیون خطی ساده (Simple Linear Regression) می‌گویند. فرم مدل رگرسیون خطی ساده به صورت زیر است:

$$Y = \beta_0 + \beta_1 x + \epsilon \quad (1)$$



شکل ۳- تغییرات NDVI در طول دوره کشت سیب زمینی

ب) شبکه‌های عصبی

$$P_{X|W}(X|I) = \sum_{k=1}^n \pi_I^k G(x, \mu_I^k, \Sigma_I^k) \quad (2)$$

در این رابطه ما سعی داریم تا احتمال پسین $P_{X|W}(X|I)$ را تخمین بزنیم. برای این کار این احتمال را به صورت ترکیب خطی از تعداد محدودی تابع گاوسی چندبعدی می‌نویسیم. پارامترهای به کاررفته در این رابطه به صورت زیر هستند:

π_I^k : ضریب (وزن) توزیع گاوسی k -ام

G : تابع گاوسی

x : ماتریس موردنظر که باید دسته‌بندی گردد.

μ_I^k : بردار میانگین توزیع گاوسی k -ام

Σ_I^k : ماتریس کوواریانس توزیع گاوسی k -ام

اجرای فرآیند، توسط الگوریتم امید ریاضی بیشینه‌سازی انجام می‌گیرد که شامل دو مرحله است:

(۱) امید ریاضی: در این مرحله یک پارامتر میانی مانند h ، روی مقادیر مرحله $k-1$ ام پارامترهای مدل رگرسیون گاوسی به دست می‌آید. به دست آوردن پارامتر h طوری انجام می‌گیرد که احتمال تخمین احتمال پسین موردنظر را حداکثر سازد. در نخستین مرحله به دلیل عدم وجود پارامترهای مرحله قبل مدل رگرسیون گاوسی مقادیر اولیه برای آن‌ها در نظر گرفته می‌شود.

(۲) مرحله بیشینه‌سازی: در این مرحله از h برای محاسبه مقادیر جدید پارامترهای مدل رگرسیون گاوسی یعنی پارامترهای مرحله k -ام استفاده می‌گردد. این دو مرحله مکرراً و پشت سر هم انجام می‌پذیرد تا تغییر پارامترهای مدل رگرسیون گاوسی به حداقل برسد و اصطلاحاً الگوریتم همگرا گردد. سپس با قرار دادن داده‌های تست به جای x احتمال قرارگرفتن داده‌های تست در هر کدام از n دسته موجود تعیین گردیده و دسته‌بندی پایان می‌پذیرد (۲۰).

د) رگرسیون بردار پشتیبان

رگرسیون بردار پشتیبان بر پایه مدل طبقه‌بندی ماشین‌های بردار پشتیبان ارائه شده است. مدل SVR برای حل مسائل رگرسیون غیرخطی استفاده می‌شود. برای توضیح مختصر SVR، فرموله سازی آن انجام می‌شود. در این راستا، مجموعه داده M که با رابطه (۲) نمایش داده شده، شامل بردارهای ورودی x_i و خروجی y_i متناظر است.

$$M = \{(x_i, y_i) | i=1, 2, \dots, n, x_i \in R^n, y_i \in R\} \quad (3)$$

که n تعداد نمونه‌های مجموعه داده را نشان می‌دهد. هدف تحلیل رگرسیون، تعیین تابع $f(x)$ به گونه‌ای است که

همان‌طور که دیده می‌شود این رابطه، معادله یک خط است که جمله خطا یا همان ϵ به آن اضافه شده است. پارامترهای این مدل خطی، عرض از مبدأ (β_0) و شیب خط (β_1) است. شیب خط در حالت رگرسیون خطی ساده، نشان می‌دهد که میزان حساسیت متغیر وابسته به متغیر مستقل چقدر است. به این معنی که با افزایش یک واحد به مقدار متغیر مستقل چه میزان متغیر وابسته تغییر خواهد کرد. عرض از مبدأ نیز بیانگر مقداری از متغیر وابسته است که به ازای مقدار متغیر مستقل برابر با صفر محاسبه می‌شود. به شکل دیگر می‌توان مقدار ثابت یا عرض از مبدأ را مقدار متوسط متغیر وابسته به ازای حذف متغیر مستقل در نظر گرفت (۱۸).

متداول‌ترین روش برای ایجاد رگرسیون غیر پارامتری و غیرخطی مبتنی بر شبکه‌های عصبی است. شبکه عصبی یک ساختار متصل از سلول‌های عصبی سازمان‌یافته در لایه است. یک نورون اساساً یک رگرسیون خطی و به دنبال آن یک تابع غیرخطی، $f(0)$ انجام می‌دهد. نورون‌های لایه‌های مختلف با پیوندهای مربوطه (وزن‌ها) به هم پیوند می‌خورند؛ بنابراین، در حالت محدود استفاده از شبکه عصبی فقط با یک نورون، نتایج مشابه (به دلیل غیرخطی بودن کمی بهتر) از نتایج به دست آمده با رگرسیون خطی خواهد داشت (۱۹).

ج) رگرسیون گاوسی

رگرسیون فرآیند گاوسی اخیراً به عنوان ابزاری قدرتمند برای رگرسیون معرفی شده است. این مدل یک رویکرد احتمالی برای یادگیری مشکلات رگرسیون عمومی با هسته‌ها فراهم می‌کند. در یک نگاه کلی، دسته‌بندی کننده‌ها به دودسته بدون ناظر و با ناظر تقسیم می‌شوند. دسته‌بندی کننده‌های بدون ناظر، برای انجام دسته بندی از ملاک‌های درونی خود برای اجرای فرآیند دسته بندی استفاده می‌کنند و در نتیجه نیازی به اطلاعات اضافی ندارند، اما در مقابل دسته‌بندی کننده‌های با ناظر علاوه بر ملاک‌های درونی برای همگرایی، نیاز به داده‌های تکمیلی نیز دارند. این اطلاعات تکمیلی، اطلاعات آموزشی است که دسته‌بندی کننده لزوماً باید روی آن آموزش ببیند.

در بین دسته‌بندی کننده‌های با ناظر، مدل رگرسیون گاوسی برای دسته‌بندی داده‌های آماری ارجحیت دارد. این مدل در ساده‌ترین حالت به صورت زیر ارائه می‌شود:

خروجی پیش‌بینی آن حداقل خطا را نسبت به خروجی مطلوب داشته باشد. تابع رگرسیون با رابطه زیر معرفی می‌شود، که δ خطای تصادفی با توزیع $(0, \sigma_2)$ است.

$$y_i = (x_i) + \delta \quad (4)$$

برای حل مسئله رگرسیون غیرخطی با SVR، نخست، ورودی‌ها به صورت غیرخطی به فضای ویژگی f با ابعاد زیاد

نگاشت می‌شود که به صورت خطی با خروجی وابستگی دارند؛ بدین منظور از رابطه زیر استفاده می‌شود.

$$(x_i) = (x_i) + b \mid w \in RN \times h, b \in R \quad (5)$$

که w بردار وزن، b مقدار بایاس و φ تابعی است که ورودی‌ها را از فضای R به فضای $h \times RN$ تصویر می‌کند (۲۱).

جدول ۲- ماتریس همبستگی مقادیر NDVI در تاریخ‌های مختلف به همراه بهره‌وری

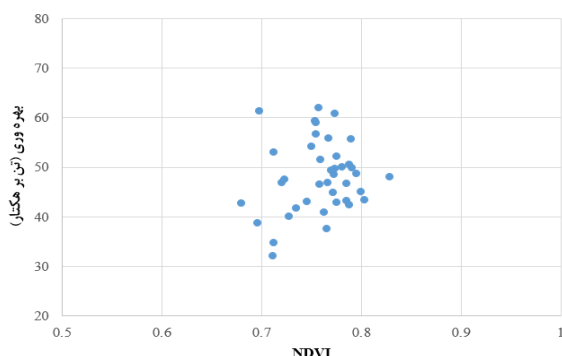
۴ مهر	۲۵ شهریور	۱۵ شهریور	۳۱ مرداد	۱۱ مرداد	۲۷ تیر	۱۲ تیر	۲۸ خرداد	۱۸ خرداد	۳ خرداد	۲۴ اردیبهشت	۳۰ فروردین	بهره‌وری
0.42	0.44	0.4	0.14	0.36	0.18	0.28	0.48	0.29	0.41	0.38	0.36	1
0.23	0.26	0.11	-0.1	-0.19	-0.5	-0.53	0.43	0.56	0.92	0.93	1	0.36
0.26	0.28	0.12	-0.13	-0.21	-0.48	-0.54	0.44	0.51	0.89	1	0.38	0.93
0.29	0.29	0.14	-0.16	-0.24	-0.44	-0.49	0.53	0.41	1	0.89	0.92	0.41
0.05	0.18	0.22	0.34	0.37	0.16	0.1	0.16	1	0.41	0.51	0.56	0.29
0.22	0.05	0.15	-0.01	0.03	0.33	0.45	1	0.16	0.53	0.44	0.43	0.48
0.28	0.2	0.34	0.63	0.65	0.94	1	0.45	0.1	-0.49	-0.54	-0.53	0.28
0.37	0.39	0.58	0.74	0.82	1	0.94	0.33	0.16	-0.44	-0.48	-0.5	0.18
0.6	0.68	0.75	0.94	1	0.82	0.65	0.03	0.37	-0.24	-0.21	-0.19	0.36
0.72	0.75	0.93	1	0.94	0.74	0.63	-0.01	0.34	-0.16	-0.13	-0.1	0.14
0.91	0.95	1	0.93	0.75	0.58	0.34	0.15	0.22	0.14	0.12	0.11	0.4
0.93	1	0.95	0.75	0.68	0.39	0.2	0.05	0.18	0.29	0.28	0.26	0.44
1	0.93	0.91	0.72	0.6	0.37	0.28	0.22	0.05	0.29	0.26	0.23	0.42

۳- بحث و نتایج

پیش‌بینی عملکرد محصولات کشاورزی از اهمیت ویژه‌ای در بازارهای داخلی و خارجی برخوردار است که زمینه‌ساز تصمیم‌های مهمی در مورد عرضه محصولات و یا انبار کردن آن‌ها می‌گردد. پایش و بررسی محصولات کشاورزی قبل از برداشت از دو جنبه اهمیت می‌یابد. جنبه اول عملکرد کشتی محصولات را مورد پژوهش قرار می‌دهد و نکاتی ارزنده برای بهبود تولید محصولات به همراه دارد. جنبه دوم اطلاعات مورد نیاز برای سازمان‌های وابسته به امور کشاورزی مانند جهاد کشاورزی و ارگان‌های وابسته به آن ارائه می‌دهد که زمینه‌ساز تصمیم‌هایی مانند کشف و

ایجاد بازارهای داخلی و خارجی برای ارائه محصولات، حذف واسطه‌ها و کمک به بهبود وضعیت کشاورزان می‌گردد. در این پژوهش برای مدلسازی داده‌های زمینی و تصاویر ماهواره‌ای و نیز تولید نقشه‌ها علاوه بر محیط متلب، از نرم افزارهای گوگل ارث انجین و QGIS استفاده شده است. داده‌های ماهواره‌ای تصاویر ماهواره سنتینل ۲ می‌باشد که از منطقه مورد مطالعه در بازه‌های زمانی مختلف اخذ گردیده است. در این تصاویر زمین‌های کشاورزی مورد مطالعه مشخص و سپس برای هر زمین شاخص گیاهی NDVI محاسبه شده است. مقادیر NDVI برای بازه زمانی کشت تا برداشت محصول سیب‌زمینی هستند. داده‌های زمینی هم شامل اطلاعاتی از میزان محصول برداشتی از هر زمین

در این مطالعه با چهار الگوریتم مختلف یادگیری ماشین اقدام به تخمین محصول سیب زمینی به منظور پایش و بررسی عملکرد محصول در بازه های زمانی مختلف از کشت تا برداشت شده است. نتایج به دست آمده از این چهار روش تخمین با سه پارامتر ضریب تعیین (R^2)، خطای میانگین مربع ریشه (RMSE) و میانگین خطای مطلق (MAE) اعتبار سنجی شدند تا مدلی انتخاب گردد که بالاترین دقت در تخمین محصول را داشته باشد (۲۲).



شکل ۴- نمودار نقطه ای NDVI در مقابل بهره وری برای تاریخ ۲۸ خردادماه

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - f_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (۶)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - f_i)^2}{n}} \quad (۷)$$

$$MAE = \frac{\sum_{i=1}^n |y_i - f_i|}{n} \quad (۸)$$

که f عملکرد پیش بینی شده، y عملکرد اندازه گیری شده، $\bar{y}$ متوسط بازده اندازه گیری شده و n تعداد بازده پیش بینی شده است.

در جدول ۳، نتایج اعتبار سنجی برای هر سه پارامتر تعیین اعتبار در چهار روش مختلف تخمین آورده شده است. در این اعتبار سنجی هرچه میزان R^2 بیشتر و میزان (RMSE) و (MAE) برای روش های تخمین کمتر باشد دقت مدل بالاتر خواهد بود. با توجه به نتایج به دست آمده الگوریتم های بردار پشتیبان و فرایند گوسی دقت بالاتری نسبت به دو الگوریتم رگرسیون خطی و شبکه عصبی دارند به عبارت دیگر الگوریتم بردار پشتیبان دقیق ترین روش برای مدل سازی سیب زمینی می باشد.

کشاورزی در طی دو سال است. با کنار هم قرار دادن داده های زمینی و ماهواره ای مدل تخمینی تولید گردید که میزان سیب زمینی برداشتی از هر زمین قبل از برداشت محصول را بتوان تخمین زد.

مطابق جدول ۲، برای تعیین میزان همبستگی داده های زمینی و ماهواره ای از ماتریسی تحت عنوان ماتریس همبستگی استفاده شده است، این ماتریس همبستگی NDVI به دست آمده از تصاویر ماهواره ای و میزان بهره وری زمین های مورد مطالعه در بازه زمانی کشت تا برداشت ارائه می دهد. رنگ سبز موجود در ماتریس همبستگی میزان وابستگی داده ها در تاریخ های مختلف را نمایش می دهد و هرچه رنگ سبز تر باشد میزان وابستگی داده ها بیشتر خواهد بود. هم چنین رنگ صورتی موجود در ماتریس همبستگی میزان وابستگی منفی یا عدم وابستگی داده ها را نمایش می دهد. با کنار هم قرار دادن داده های زمینی و ماهواره ای مدل تخمینی تولید گردید که میزان سیب زمینی برداشتی از هر زمین قبل از برداشت محصول را بتوان تخمین زد. برای مدل سازی از روش های تک متغیره استفاده شد به این دلیل که با توجه به ماتریس همبستگی در بازه های زمانی میزان وابستگی بین داده زمینی و NDVI بسیار زیاد بود که به رنگ سبز نمایش داده شده است. اگر برای مدل سازی از روش های چند متغیره استفاده می شد به دلیل وابسته بودن بازه های زمانی مختلف با اطلاعات تکراری مواجه می شدیم که تکراری بودن اطلاعات بی فایده می باشد؛ بنابراین از روش تک متغیره برای مدل سازی سیب زمینی استفاده گردید.

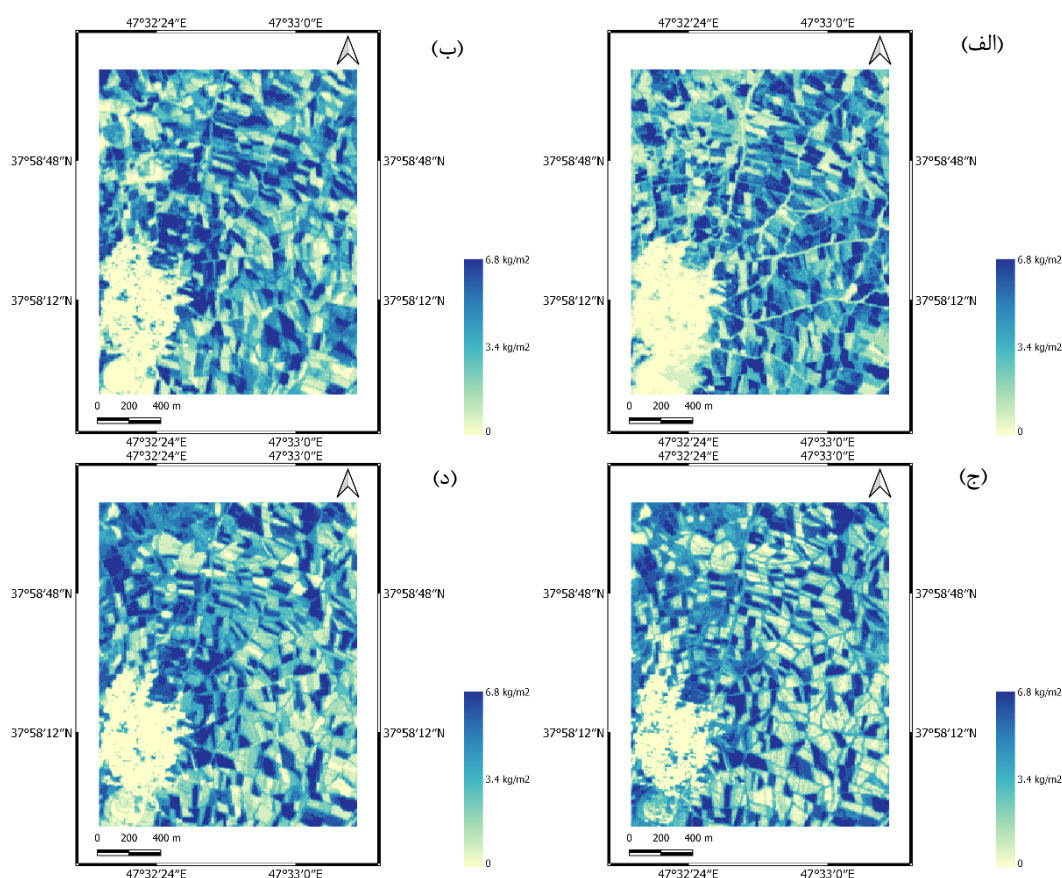
با توجه به ماتریس همبستگی بیشترین مقدار همبستگی NDVI و بهره وری در تاریخ ۲۸ خردادماه اتفاق می افتد که میزان این همبستگی ۰٫۴۸ است که همبستگی پایینی در مقایسه با غلات دارد. میزان همبستگی در غلات عددی نزدیک به یک است چون NDVI بیشتر روی کلروفیل و سبزیگی گیاهان تعیین می گردد و سبزیگی هم ارتباط مستقیم با میزان غلات تولیدی دارد ولی سیب زمینی به صورت غده کشت می یابد و سبزیگی تأثیری به اندازه غلات در بهره وری آن ندارد. برای مدل سازی تمام همبستگی هایی که بالای ۰٫۳ قرار داشتند از روش تک متغیره استفاده شد و بهترین مدل ها برای تاریخ ۲۸ خردادماه که بیشترین میزان همبستگی را داشت تولید شدند (شکل ۴).

جدول ۳- نتایج اعتبار سنجی مدل‌های تخمین

روش تخمین	R ²	RMSE (تن بر هکتار)	MAE (تن بر هکتار)
رگرسیون خطی	۰,۴۲	۸,۳۱	۷,۲۹
شبکه عصبی	۰,۵۰	۹,۰۲	۸,۴۴
بردار پشتیبان	۰,۵۸	۶,۸۸	۷,۱۵
فرآیند گوسی	۰,۶۱	۷,۴۵	۷,۳۸

در شکل ۵، نقشه به‌دست‌آمده از چهار روش تخمین: الف) رگرسیون خطی، ب) شبکه عصبی، ج) بردار پشتیبان، د) فرآیند گوسی نمایش داده شده است. این نقشه‌ها نمایش‌دهنده میزان محصول برداشتی برحسب کیلوگرم بر مترمربع هستند که بیشترین مقدار آن ۶,۸ می‌باشد. با توجه

به نقشه‌ها زمین‌های کشاورزی که از نظر آبیاری مشکل نداشتند و آبیاری آن‌ها در بازه زمانی مشخص انجام گرفته بود محصول بهتری تولید کرده بودند. از عوامل مؤثر دیگر در تولید محصول بهتر از نظر کیفیت و کمیت می‌توان به کود دهی مناسب و سم‌پاشی به‌موقع و تدابیری که کشاورزان بر اساس تجربه اتخاذ می‌کنند، اشاره کرد. این زمین‌ها دارای محصول پربارتری بودند که در تصاویر به رنگ آبی تیره نمایش داده شدند. به‌طور کلی عواملی نظیر دما، بارش، آبیاری، کود دهی، سم‌پاشی و کشت صحیح سیب-زمینی تأثیر مستقیم در کشت محصول خواهند داشت.



شکل ۵- نقشه به‌دست‌آمده از چهار روش تخمین: الف) رگرسیون خطی، ب) شبکه عصبی، ج) بردار پشتیبان، د) فرآیند گوسی

۴- نتیجه گیری

در این مقاله با استفاده از تصاویر ماهواره‌ای سنتینل-۲ و یک مدل رگرسیون، محصول سیب زمینی ۴۰ زمین کشاورزی در شهرستان سراب، استان آذربایجان شرقی، توسعه یافت. شاخص گیاهی NDVI از تصاویر ماهواره‌ای استخراج شدند و میزان همبستگی آن‌ها با داده‌های زمینی (میزان بهره وری هر مزرعه) به دست آمد. بیشترین

همبستگی در تاریخ ۲۸ خرداد به میزان ۰,۴۸ بود. دلیل پایین بودن این همبستگی غده بودن سیب‌زمینی و تأثیر کم سبزی‌نگی روی محصول تولیدی می‌باشد. از روش تک متغیره برای مدل‌سازی استفاده شد و تاریخ ۲۸ خرداد به‌عنوان بازه مدل‌سازی انتخاب گردید. چون با استفاده از روش‌های چند متغیره با توجه به این‌که همبستگی زیادی در اطلاعات بازه‌های زمانی مختلف وجود داشت اطلاعات

وضعیت سلامت آن را پیش‌بینی کرد. نتایج این تحقیق نشان داد که مدل‌های تخمین، اطلاعات بسیار ارزنده‌ای به منظور مدیریت بهتر محصولات زراعی به‌همراه خواهد داشت تا از هدر رفت منابع و جلوگیری از واردات مورد استفاده برنامه ریزان حوزه کشاورزی قرار گیرند.

تکراری زیادی تولید می‌شد و بهمین دلیل از روش تک متغیره استفاده کردیم. برای مدل‌سازی از چهار الگوریتم یادگیری ماشین استفاده شد که روش‌های گوسی و بردار پشتیبان بهترین نتایج را رقم زدند. با توجه به مدل تخمین تولیدشده می‌توان سه ماه قبل از برداشت محصول با دقت ۵۰ درصد میزان سیب‌زمینی تولیدی هر زمین کشاورزی و

مراجع

- [1] Lotze-Campen, H.; Müller, C.; Bondeau, A.; Rost, S.; Popp, A.; Lucht, W. Global food demand, productivity growth, and the scarcity of land and water resources: A spatially explicit mathematical programming approach. *Agric. Econ.* 2008, 39, 325–338.
- [2] Windfuhr, M.; Jonsén, J. Food Sovereignty: Towards Democracy in Localized Food Systems. 2005.
- [3] Chen, F.; Kissel, D.E.; West, L.T.; Adkins, W. Field-scale mapping of surface soil organic carbon using remotely sensed imagery. *Soil Sci. Soc. Am. J.* 2000, 64, 746–753.
- [4] Gebbers, R.; Adamchuk, V.I. Precision agriculture and food security. *Science* 2010, 327, 828–831.
- [5] Conijn, J.G.; Bindraban, P.S.; Schröder, J.J.; Jongschaap, R.E.E. Can our global food system meet food demand within planetary boundaries? *Agric. Ecosyst. Environ.* 2019, 251, 244–256.
- [6] Al-Gaadi KA, Hassaballa AA, Tola E, Kayad AG, MadugunduR, Alblewi B, et al. (2016) Prediction of Potato Crop Yield Using Precision Agriculture Techniques. *PloS ONE* 11(9).
- [7] Seelan, S.K.; Laguet, S.; Casady, G.M.; Seielstad, G.A. Remote sensing applications for precision agriculture: A learning community approach. *Remote Sens. Environ.* 2003, 88, 157–169.
- [8] Pathak, B.T. 2005. Validation of an existing potato yield model using airborne multispectral remote sensing data. Unpublished MS thesis. Utah State University, Logan, Utah.
- [9] Timlin, D.; Rahman, L.; Baker, S.M.; Reddy, J.; Fleisher, D.; Quebedeaux, B. Whole plant photosynthesis, development, and carbon partitioning in potato as a function of temperature. *Agron. J.* 2006, 98, 1195–1203.
- [10] Bala, S.K., and A.S. Islam. 2009. Correlation between potato yield and MODIS derived vegetation indices. *International Journal of Remote Sensing* 30:10, 2491- 2507.
- [11] Devaux, A.; Kromann, P.; Ortiz, O. Potatoes for Sustainable Global Food Security. *Potato Res.* 2014, 57, 185–199.
- [12] Haverkorta, A.J.; Struik, P.C. Yield levels of potato crops: Recent achievements and future prospects. *Field Crop. Res.* 2015, 182, 76–85.
- [13] Zhao, J.; Zhang, Y.; Qian, Y.; Pan, Z.; Zhu, Y.; Zhang, Y.; Guo, J.; Xu, L. Coincidence of variation in potato yield and climate in northern China. *Sci. Total Environ.* 2016, 573, 965–973.
- [14] Asadian, Gh., Azimnejad, Z., Shojaei, H., (2017). "Investigation of 9-years performance of saffron in climatic conditions of Hamedan province". Sixth Scientific Research Congress on the Development and Promotion of Agriculture and Natural Resources in Iran.
- [15] Alizadeh-Salteh, S., Nayyerpoor, A., Hooshmand, S., (2017). "Evaluation of Socio-economic effects and profit efficiency of saffron cultivation in Bonab of Marand". *Journal of Saffron Research (semi-annual)*. Vol.5, No.1, p. 78-89.
- [16] Parhizgar, A., Ghaffari Gylandeh, (2006). "GIS and multi criteria decision analysis". The publication samt ,Tehran.
- [17] Dong, X., Vuran, M.C., Irmak, S., 2013. Autonomous precision agriculture through integration of wireless underground sensor networks with center pivot irrigation systems. *Ad Hoc Networks*, 11:1975-1987.
- [18] Wright, J.L., and J.C. Stark. 1990. Potato. Irrigation of agricultural crops. *Agronomy Monograph* No. 30. ASA-CSSA-SSSA, Madison, WI. 30.
- [19] Redulla, C.A. 2002. Relating potato yield and quality to field scale variability in soil characteristics. *American Journal of Potato Research*, Sep/Oct 2002.

- [20] Jayanthi, H., C.M.U. Neale, and J.L. Wright. 2007. Development and validation of canopy reflectance-based crop coefficient for potato. *Agricultural water Management* 88:235-246.
- [21] Jayanthi, H. 2003. Airborne and Ground-Based Remote sensing for the estimation of Evapotranspiration and yield of Bean, Potato, and sugar beet crops. Unpublished Ph.D dissertation. Utah State University, Logan, Utah. 185 p.
- [22] Newton, I.H.; Islam, T.; Saiful, I. Yield Prediction Model for Potato Using Landsat Time Series Images Driven Vegetation Indices. *Remote Sens. Earth Syst. Sci.* 2018, 1, 29–38.