

## مدل سازی وضعیت صخره های مرجانی با استفاده از رگرسیون ماشین بردار پشتیبان و اعمال شاخص های طیفی

صادق اسلامی<sup>۱</sup>، مهدی حسنلو<sup>۲</sup>

s.eslami@ut.ac.ir

۱- دانشجوی کارشناسی ارشد دانشکده مهندسی عمران، دانشگاه تهران  
۲- استادیار دانشکده مهندسی نقشه برداری، دانشگاه تهران

### چکیده

از منظر تنوع زیستی، جزایر مرجانی مانند جنگل های استوایی، متنوع ترین اکوسیستم های جهان و شاخصی از سلامت اکوسیستم محسوب می شوند اما اکثر این جزایر به دلیل تحولات توریستی و نیز تأثیرات تغییر آب و هوا رو به نابودی هستند. توانایی شناسایی و ارزیابی سلامت مرجان ها با استفاده از تصاویر ماهواره ای شیوه ای مقرون به صرفه و موثر است. با توجه به اثرگذاری دمای سطح آب بر روی سلامت و توزیع مرجان ها، با استفاده از الگوریتم هایی به رابطه بین آن ها و سلامت مرجان های منطقه مورد مطالعه که در این مقاله جزیره مرجانی هرون استرالیا است، پرداخته شده است. در این مطالعه با استفاده از تصاویر لندست-۸ و به دست آوردن بازتابش باندها، ویژگی ها<sup>۱</sup> و شاخص های طیفی مهم مرتبط با آب و صخره های مرجانی مانند: FAI, AWEI, SWI, MNDWI, GRVI, NDWI, NDVI, در این مطالعه با استفاده از روش رگرسیون بردار پشتیبان (SVR) و بدست آوردن پارامترهای کرنل آن، به مدل سازی وضعیت سلامت صخره های مرجانی پرداخته شده است. در این مقاله از الگوریتم های تکاملی مثل الگوریتم ژنتیک (GA) و الگوریتم انتخاب ویژگی ترتیبی جلویی (SFS) برای رسیدن به انتخاب ویژگی مطلوب و مدل سازی بهینه و طبقه بندی سلامت مرجان ها استفاده شده است. در این مقاله با استفاده از روش رگرسیون بردار پشتیبان در حالت کلی به  $R^2=0.979$  و  $RMSE=0.591$  و در حالت پیشنهادی (SVR-GA) به  $R^2=0.983$  و  $RMSE=0.53$  رسیدیم که بیانگر عملکرد خوب این مدل بهینه است.

واژگان کلیدی: شاخص های طیفی، وضعیت سلامت، SVR، جزیره هرون استرالیا، تصاویر لندست-۸، انتخاب ویژگی.

تاریخ دریافت مقاله : ۹۷/۰۵/۱۰  
تاریخ پذیرش مقاله : ۹۷/۰۹/۰۲

<sup>۱</sup> Feature

۱- مقدمه

در میان اقیانوس‌های جهان یکی از غنی‌ترین فرم‌های تنوع زیستی را در ساختارهای زنده‌ای می‌یابیم که مرجان نام دارند. آبنگ‌های مرجانی که به‌ظاهر شبیه صخره‌هایی بی‌جان می‌باشند، در واقع جوامع زنده‌ای هستند که تنوع بالای موجودات ساکن آنها را در کمتر اکوسیستمی می‌توان شاهد بود. از منظر تنوع زیستی، جزایر مرجانی مانند جنگل‌های استوایی، متنوع‌ترین اکوسیستم‌های جهان هستند. این آبیان حساس به‌ویژه در مناطق گرمسیری جهان نقش بسیار مهمی را در تداوم حیات در اقیانوس‌ها ایفا می‌کنند. همچنین آبنگ‌های مرجانی نقش مهمی در پایداری جوامع محلی دارند [۱].

همچنین شرایط محیطی نیز در شکل‌گیری صخره‌ها دخالت دارد. مرجان‌ها به آب‌های ۲۲ تا ۳۲ درجه سانتی‌گراد احتیاج دارند و از سطح تا اعماق ۳۰ متر، میان عرض جغرافیایی ۲۶ درجه شمالی و ۲۶ درجه جنوبی به سر می‌برند. آبنگ‌های مرجانی به‌عنوان مکان تغذیه‌ای، تولیدمثلی، پرورش نوزادان و نیز پناهگاه ماهیان و بسیاری از موجودات آبی عمل می‌کنند و نیز شاخصی از سلامت اکوسیستم هستند. بسیاری از موجودات نیز به‌صورت همزیست با مرجان‌ها زندگی می‌کنند. از این‌رو تراکم و پراکنش مرجان‌ها باعث تغییر در جمعیت ماهی‌ها و دیگر موجودات می‌شود. از موارد کاربرد سنجش‌ازدور در مدیریت مناطق ساحلی، حفظ محیط‌زیست و منابع طبیعی و بررسی عارضه‌ها در این مناطق است. علم سنجش‌ازدور با توجه به قابلیت پایش مستمر و مداوم و قابلیت ذخیره‌سازی برای مقاصد بلندمدت و توانایی بررسی اکوسیستم بدون تماس مستقیم با آن و همچنین صرفه‌جویی در وقت و در زمان می‌تواند برای بررسی صخره‌ها مفید باشد. صخره‌های مرجانی به‌عنوان مهم‌ترین زیستگاه‌های جانوری در دریا، همواره بیش‌ترین آسیب را از فعالیت‌های انسانی در مناطق دریایی متحمل شده‌اند و وجود آن‌ها به لحاظ مسائل سلامت اکوسیستم و محیط‌زیست و اکولوژی و همچنین فعالیت‌های اقتصادی و تجاری نظیر جذب توریست حائز اهمیت است [۴]. آب دریا باید روشن و شفاف باشد تا اطمینان حاصل شود که عبور نور صورت می‌گیرد. این حساسیت مرجان‌ها که تحت شرایط خاصی به وجود می‌آید این امکان را به ما می‌دهد که ما با استفاده از سنجش‌ازدور

با استفاده از باندهای مرئی و غیرمرئی مرجان‌ها را بررسی کنیم [۳ و ۲]. از اولین پژوهش‌های انجام‌گرفته در این زمینه می‌توان به ۱۹۷۰ اشاره کرد که از داده‌های لندست برای تصاویر رقومی مرجان‌های بزرگ استرالیا و ژئومورفولوژی آن استفاده شد که منجر به طبقه‌بندی اکوسیستم مرجانی با تصویر سنجنده TM با جزئیات کم شامل چهار کلاس مرجان، جلبک، ماسه و علف دریایی با دقت‌های ۸۶٪ و ۴۷٪ و ۸۳٪ و ۵۹٪ بود [۱ و ۴]. در سال ۲۰۰۰ با استفاده از داده‌های ابرطیفی هوایی که توان تفکیک مکانی و طیفی بالایی دارند جوامع زیستی مختلف اعماق خلیجی در هاوایی با استفاده از انتخاب طول‌موج مناسب و توابع خطی به گروه‌های مرجان‌ها و جلبک‌ها و شن و ماسه طبقه‌بندی شد [۵]. در سال ۲۰۰۳ ده تصویر IKONOS از مکان‌های مختلف توزیع صخره‌های مرجانی در سراسر جهان برای ارزیابی پتانسیل توان تفکیک ۴ متری برای نقشه‌برداری زیستگاه صخره‌های مرجانی پردازش شد. پردازش شامل تصحیحات زبری سطح دریا و عمق سنجی بود و نتایج حاصل شامل طبقه‌بندی زیستگاه‌های صخره ساده متوسط تا پیچیده از ۵ تا ۱۱ کلاس بود. پس از آن به مقایسه دقت طبقه‌بندی IKONOS با طبقه‌بندی تصاویر ماهواره لندست-۷ پرداختند [۶]. دسته دیگر مطالعات در زمینه بهینه‌سازی سیستم‌های طبقه‌بندی، به مسئله انتخاب زیرمجموعه بهینه باندهای ورودی می‌پردازند. ژو و همکاران در سال ۲۰۰۸ الگوریتم ارائه‌شده در را بر روی تصاویر فراطیفی هایپریون آزمایش کردند و مشاهده گردید نتایج به‌دست‌آمده نسبت به زمانی که از همه باندها استفاده‌شده باشد، بهبود یافته است. در این مطالعه تأثیر روش پیشنهادی تنها با استفاده از یک هسته و یک روش چند کلاسه در ماشین‌های بردار پشتیبان بررسی شد [۷]. جونگ و همکاران در سال ۲۰۱۲ در شهر وهان<sup>۱</sup> چین به مقایسه بین دو طبقه‌بندی شبکه عصبی و ماشین بردار پشتیبان بر روی تصاویر ETM+ به منظور استخراج کاربری اراضی پرداختند که نتایج بیانگر دقت بالای ۹۷٪ برای هر دو نوع طبقه‌بندی بود (بیشتر بودن دقت کلی ۱،۴ درصدی روش ماشین بردار پشتیبان نسبت به شبکه عصبی) [۱].

<sup>۱</sup> Wuhan

بر اثر گرمای بیش از اندازه درنهایت منجر به سفیدی مرجان‌ها (مردن آن‌ها) می‌شود. پس با اندازه دمای سطح آب و تولید نقشه‌های دمای سطحی آب در بازه‌های مختلف به سلامت و اکوسیستم مرجان‌ها در تمام نقاط جهان پی می‌بریم. الگوریتم Hotspots Coral Bleaching با استفاده از  $SST_{(daily)}$  دمای سطح روزانه و  $MMM^2$  بیشترین متوسط ماهانه) و رابطه (۱) به صورت روزانه نقشه‌ای را تولید می‌کند [۹].

$$HS = \begin{cases} SST_{daily} - MMM & SST_{daily} > MMM \\ 0 & SST_{daily} \leq MMM \end{cases} \quad (1)$$

با استفاده از الگوریتم‌های HOTSPOTS و بازه‌های تعریف شده برای آن نقشه‌هایی به منظور بررسی سلامت محیط‌زیست مرجان‌ها تهیه می‌شود (جدول (۱)) در این مقاله ابتدا داده‌های دمای سطحی آب (SST) مربوط به سنجنده مدیس برای منطقه مورد نظر جمع‌آوری می‌شود. سپس با استفاده از تصاویر لندست-۸ بازتابش باندها را به دست می‌آوریم و به کمک شاخص‌های طیفی مرتبط با صخره‌های مرجانی و کمیت‌هایی مثل دمای درخشندگی و رنگ دریا که از لندست-۸ استخراج می‌شوند، با استفاده از روش SVR (رگرسیون بردار پشتیبان) به مدل‌سازی و طبقه‌بندی وضعیت مرجان‌ها می‌پردازیم و سپس به کمک الگوریتم‌های تکاملی مثل ژنتیک و SFS<sup>۳</sup> به بررسی و انتخاب ویژگی‌ها و مدل بهینه می‌پردازیم [۱۱].

جدول (۱) الگوریتم جهانی نظارت و نمایش صخره‌های مرجانی.

تعریف	میزان استرس
$HotSpot \leq 0$	بدون استرس (بی خطر)
$0 < HotSpot < 1$	در آستانه سفید شدگی مرجان‌ها
$1 \leq HotSpot$	هشدار (دارای خطر)

## ۲-۲- شاخص‌های طیفی

اساس استفاده از سنجش‌ازدور، با شناسایی و آشکار کردن تغییرات پوشش زمین با استفاده از تغییر در مقدار بازتابش ممکن می‌شود. از این رو شاخص‌ها و روش‌های مختلف از مقدار بازتابش و بازتابندگی عوارض مختلف برای شناسایی،

مقیاس انجام گرفت که نتایج شامل ایجاد شش کلاس (مرجان نرم، مرجان سخت، جلبک، علف دریایی، مرجان مرده، ما سه) با دقت بالای ۷۰٪ بود [۱]. در این مقاله برای اولین بار با توجه به اهمیتی که صخره‌های مرجانی برای حفظ محیط‌زیست دارند و ذکر این نکته که آسیب دیدن یا نابودی آن‌ها، به قیمت از بین رفتن گونه‌های دریایی زیادی است که می‌تواند در گذر زمان اثرات مخربی را بر محیط زیست دریایی داشته باشد، با کمک گرفتن از الگوریتم نظارت بر صخره‌های مرجانی (NOAA) به مدل‌سازی وضعیت سلامت مرجان‌های جزیره هرون استرالیا به روش رگرسیون ماشین پشتیبان بردار (SVR) پرداخته ایم [۱]. در این مطالعه در مورد ویژگی‌ها و شاخص‌های متعددی که مرتبط با آب‌ها و صخره‌های مرجانی است صحبت شده است. این ویژگی‌ها شامل بازتابش باندها (۷ باندها)، دمای درخشندگی ۲ باندها لندست-۸ (باندهای ۱۰ و ۱۱)، ۷ شاخص FAI, AWEL, SWI, NDWI, NDVI, MNDWI, GRVI و کمیت رنگ دریا است که با استفاده از الگوریتم‌های انتخاب ویژگی مثل ژنتیک و انتخاب جلویی ترتیبی، به بهترین ویژگی‌ها برای مسئله مربوطه رسیده‌ایم [۸].

در مسئله مربوط به SVR با چالش انتخاب مناسب پارامترهای کرنل گوسی شامل C و  $\epsilon$  و  $\gamma$  برای مدل خود و ارزیابی الگوریتم‌های انتخاب ویژگی و ترکیب آن‌ها هستیم و پاسخ به این سوال که آیا استفاده از الگوریتم‌های انتخاب ویژگی و مدل‌های پیشنهادی می‌تواند بهتر از روش کلی SVR برای مدل‌سازی سلامت صخره‌های مرجانی منطقه مورد مطالعه باشد؟

## ۲- روش‌های اجرایی

### ۱-۲- الگوریتم نظارت بر صخره‌های مرجانی (NOAA)

یکی از الگوریتم‌های استاندارد مورد استفاده در جهان الگوریتم‌های مربوط الگوریتم اداره نظارت و نمایش صخره‌های مرجانی (NOAA) است که با استفاده از تولید نقشه‌ها و الگوریتم‌های مختلف و همچنین رابطه بین گرما و دمای آب سلامت مرجان‌ها را طبقه‌بندی می‌کند. ایجاد شدن تنش‌های حرارتی و اختلاف بین دمای سطحی آب روزانه و دمای سطحی متوسط منطقه (۱۵ روز تا ۳۰ روزه)

<sup>2</sup> Maximum of Monthly Mean

<sup>3</sup> Sequential Feature Selection

<sup>1</sup> National Oceanic and Atmospheric Administration

کشاورزی را می‌توان بهبود بخشید. این شاخص ماهواره‌ای با استفاده از انعکاس در باند فرورسرخ نزدیک (NIR) و باند سبز (G) و ترکیب آن به دست می‌آید. به نوعی می‌توان گفت شاخصی است از میزان آبی که آن منطقه از دست داده که می‌تواند منجر به استرس برای گیاهان دریایی شود. این شاخص عددی است بین +۱ و -۱. اگر این شاخص بین ۰ و -۱ باشد به معنای کم بودن آب مربوط به گیاهان و کمبود گیاهان به دلیل کمبود آب است و اکثراً شامل مناطق صیقلی و صاف زمین می‌باشند و هرچه شاخص به +۱ نزدیک باشد به معنی بیشتر بودن آب و کم شدن استرس برای گیاهان مطابق رابطه (۲) است [۱۲].

### ۲-۲-۲- شاخص آب تفاضلی نرمال اصلاحی (MNDWI)

مشابه و اصلاح شده NDWI است می‌تواند منجر به افزایش شناسایی ویژگی‌های آب، به دلیل از بین بردن مؤثر اختلال‌های ناشی از زمین و پوشش گیاهی و خاک در اقیانوس‌ها شود اما به جای انعکاس باند فرورسرخ نزدیک در لندست-۸ از باند فرورسرخ کوتاه استفاده می‌شود. بر این اساس MNDWI برای افزایش و استخراج اطلاعات آب برای یک منطقه آب مناسب‌تر از NDWI است پس اگر می‌خواهیم بر روی مناطق ساحلی و مرزی کار کنیم بهتر است از MNDWI استفاده کنیم [۱۳]. علاوه بر این، MNDWI می‌تواند جزئیات بیشتری از آب را نشان می‌دهد و برای تشخیص تفاوت‌های ظریف در کیفیت آب مفید و مطابق رابطه (۳) است:

$$MNDWI = \frac{GREEN - SWIR}{GREEN + SWIR} \quad (3)$$

### ۲-۲-۳- شاخص گیاهی سبز قرمز (GRVI)

استفاده از شاخص پوشش گیاهی سبز قرمز (GRVI) می‌تواند به عنوان یک ارزیابی شاخص رشد بر اساس مشاهدات در سطح پایه با استفاده از انعکاس طیفی و فنولوژی به منظور بررسی اکوسیستم منطقه باشد که همواره عددی بین ۰ و ۱۰ است. از نظر تعادل بین بازتاب در باند سبز (طول موج حدود ۵۰۰-۵۷۰ نانومتر) و بازتاب در باند قرمز (طول موج از حدود ۶۲۰-۷۰۰ نانومتر) ما می‌توانیم سه گروه از الگوهای بازتاب طیفی برای اجزای اصلی پوشش زمین پیدا کنیم در پوشش گیاهی سبز (مخروطیان،

در باندهای مختلف طیفی الکترومغناطیس استفاده می‌کنند. در زمینه شناسایی سطح آب ماهواره‌های زیادی قابل استفاده هستند که دارای توان تفکیک مکانی، زمانی و طیفی گوناگون هستند. اولین ماهواره‌ای که از آن برای تصویر کردن پهنه‌های آب و سیلاب‌ها مورد استفاده قرار گرفت، سری ماهواره‌های لندست و با استفاده از سنجنده MSS بود، که تصویربرداری خود را از سال ۱۹۷۲ آغاز کرده بود و از جمله سنجنده‌های اپتیکی است. به طور کلی تصاویری برای شناسایی سطوح آبی مناسب هستند، که دارای باند در طول موج‌های جذبی و بازتابی آب در محدوده طیف الکترومغناطیس باشند [۱۲]. از میان ماهواره‌های مختلف، ماهواره لندست از جمله ماهواره‌های پرکاربرد برای شناسایی پارامترهای آب و شاخص‌های آن است. شاخص‌ها با استفاده از ترکیب باندهای مختلف و یا استفاده از تضاد محدوده طیفی باندها به بررسی پدیده‌ها و ویژگی‌های مختلف می‌پردازند و در واقع باعث بهبود در سیگنال طیفی دریافتی، شناسایی بهتر و خارج شدن نویز از قسمت‌های مختلف طول موج‌ها می‌شوند [۱۲]. در ادامه به بررسی و ترکیب‌های این شاخص‌ها پرداخته‌ایم. فقط این نکته قابل ذکر است که در تمامی رابطه‌ها به جای استفاده از انعکاس باندها از عنوان‌های باندها استفاده شده است و استخراج آن‌ها توسط محققین با استفاده از تصاویر ماهواره‌ای بدست آمده است.

### ۲-۲-۱- شاخص آب تفاضلی نرمال شده (NDWI)

این شاخص در واقع شاخص آب تفاضلی نرمال شده فیتزرا<sup>۱</sup> است و در واقع حالت جدید و اصلاح شده NDWI است که در آن میزان روشنایی پهنه‌های آبی نسبت به شاخص NDWI بیشتر بوده و پوشش گیاهی از روشنایی بسیار کم و خاک از درجات خاکستری متوسطی برخوردار است. در طول خشک‌سالی، حجمی از پوشش گیاهی توسط آب در استرس هستند. این می‌تواند تأثیر عمده‌ای بر رشد گیاه به طور کلی داشته و می‌تواند منجر به شکست محصول و یا تولید محصول کمتر در مناطق کشاورزی شود. تشخیص زودهنگام تنش آب گیاه می‌تواند برای جلوگیری از چنین عواقبی مفید باشد. با ارائه اطلاعات نزدیک به واقعی بر استرس آب گیاه می‌توان غلبه کرد و سطح مدیریت آب و

<sup>1</sup> Feeters

فروسرخ نزدیک و فروسرخ کوتاه جذب نور بیشتری انجام می‌دهد. در واقع، با توجه به این‌که در این باندها جذب بالاست، آب کدر در محیط به رنگ سیاه در می‌آید. با توجه به باندهای مهمی که در مورد جذب آن صحبت شد و استفاده از پایه خطی به شاخصی به نام جلبک شناور تعریف می‌شود [۱۵].

$$R_{rc} = (3.1415L/F_0 \cos\theta_0) - R_r \quad (5)$$

L در رابطه (۵) بیانگر بازتابش فروسرخ نزدیک و R بیانگر انعکاس این باند است.  $F_0$  تابش فرازمینی خورشیدی<sup>۲</sup> است که با توجه به ماه و روز و عرض جغرافیایی تعیین می‌شود.  $\theta_0$  زاویه زنیته خورشیدی است که در فایل مرتبط با تصاویر لندست-۸ موجود است. همچنین در رابطه ۶،  $\lambda$  اندازه طول موج است که برای باند فروسرخ نزدیک ۸۵۹، برای باند فروسرخ کوتاه ۱۲۴۰ و برای باند قرمز ۶۴۵ نانومتر است.

$$R'_{rc,NIR} = R_{rc,RED} + (R_{rc,SWIR} - R_{rc,RED}) \times (\lambda_{NIR} - \lambda_{RED}) / (\lambda_{SWIR} - \lambda_{RED}) \quad (6)$$

شاخص جلبک‌های شناور از تفاضل دو رابطه (۵) و (۶) بدست می‌آید.

$$FAI = R_{rc,NIR} - R'_{rc,NIR} \quad (7)$$

#### ۲-۲-۵- شاخص رطوبت سطح (SWI)

شاخص رطوبت سطح (SWI) برای شناسایی پهنه‌های آبی استفاده می‌گردد. مشخصه‌های طیفی آب به‌طور مشخص با بقیه عوارض متفاوت است و آب زلال دارای بازتابندگی کم در ناحیه طیفی مرئی است.

با توجه به بررسی‌های گوناگون در زمینه شناسایی سطح آب با استفاده از سنجش‌ازدور ماهواره‌ای توسط محققین و بررسی شاخص‌ها و باندهای مختلف الکترومغناطیس، محدوده طیفی فروسرخ محدوده مناسبی برای شناسایی پهنه‌های آبی است. شاخص SWI نیز از جمله شاخص‌های شناسایی سطح آب است. این شاخص برای پهنه‌های آبی دارای بیشترین مقدار و برای زمین‌های خشک دارای کمترین مقدار است.

البته لازم به ذکر است که این شاخص بیشتر برای نشان دادن پیکسل آب است، که این پیکسل آب می‌تواند آب داخل خاک عمیق در جنگل‌های متراکم نیز باشد و رابطه آن مطابق (۸) است [۱۵].

درختان برگریز، و چمن) بازتاب طیفی سبز بالاتر از قرمز است. خاک قهوه‌ای (شن و ماسه، گل‌ولای و خاک رس خشک) بازتاب طیفی سبز کمتر از قرمز است و بازتاب طیفی آب و برف شبیه هم هستند. مطالعات نشان داده که روابطی بین مقادیر GRVI و تغییرات فصلی پوشش گیاهی و سطح زمین با توان تفکیک زمانی بالا برقرار است. هر چه این شاخص بیشتر باشد تراکم بیشتر و سبز بودن برگ را نشان می‌دهد. مقدار کم این شاخص بیانگر تراکم کم برگ‌ها و تغییر رنگ ظاهری و رو به پاییز رفتن منطقه است. پس بررسی این شاخص برای بررسی اقلیم و اکوسیستم منطقه مفید باشد [۱۴] که به صورت رابطه (۴) است:

$$GRVI = \frac{GREEN - RED}{GREEN + RED} \quad (4)$$

#### ۲-۲-۴- شاخص جلبک شناور (FAI)

انواع جلبک‌های شناور در اقیانوس‌های باز و آب‌های ساحلی یافت می‌شود، اما تشخیص این ویژگی با استفاده از داده‌های ماهواره‌ای سنتی و الگوریتم به علت عدم تفکیک فضایی و پوششی و فرکانس مجدد، یا به دلیل ذاتی دشوار و یا حتی غیرممکن است. در اینجا، یک شاخص رنگ اقیانوس ساده یعنی شاخص جلبک شناور مورد استفاده برای شناسایی جلبک‌های شناور در محیط‌های باز از اقیانوس‌ها با استفاده از توان تفکیک مکانی متوسط (۲۵۰ تا ۵۰۰ متر اسپکترورادیومتر تصویربرداری) که بیشتر توسط مدیس اندازه‌گیری می‌شود، را توسط لندست-۸ بررسی می‌کنیم. این شاخص با استفاده از تفاوت بین بازتاب در طول موج ۸۵۹ نانومتر (مربوط به پوشش گیاهی لبه قرمز) و یک پایه خطی<sup>۱</sup> بین باند قرمز (۶۴۵ نانومتر) و فروسرخ کوتاه (۱۲۴۰ یا ۱۶۴۰ نانومتر) تعریف می‌شود. با توجه به مطالعات انجام‌شده این شاخص به نسبت شاخص‌های سنتی موجود مثل NDVI با توجه به موقعیت خود حساسیت کمتری نسبت به انواع گرد و خاک و موقعیت خورشید و سان گلینت دارد و همچنین در مواقع ابرهای نازک هم قابل محاسبه است [۱۵].

از سوی دیگر، جلبک شناور در آب در باند فروسرخ نزدیک سطح بازتاب بالاتری نسبت به طول‌موج‌های دیگر دارد و می‌تواند به راحتی از آب‌های اطراف متمایز است. در مقایسه با سطوح زمین، آب به شدت در طول‌موج‌های سرخ،

<sup>2</sup> Extraterrestrial Solar Radiation

<sup>1</sup> Linear Baseline

کاهش اثرات اتمسفری، توپوگرافی و همچنین تصحیح تغییرات تابش ورودی خورشید، از قابلیت‌های این شاخص بوده ولی تداخل بین خاک و گیاه در محدوده فرسوخ را برطرف نمی‌کند. از طرفی چون اختلاف بازتاب پوشش گیاهی در این دو باند بیشتر از اختلاف بازتاب خاک در آن‌ها است، خاک بسیار کمتر از پوشش گیاهی تحت تأثیر این فرآیند قرار می‌گیرد [۱۵]. بازتابش طیفی گیاهان سالم (یا کلروفیل) بیشتر در طول موج الکترومغناطیسی فرسوخ نزدیک (NIR) و طول موج سبز نسبت به سایر طول‌موج‌ها اتفاق می‌افتد. گیاهان سالم که حاوی کلروفیل هستند، نور آبی و بیشتر طول‌موج‌های قرمز را جذب می‌کنند. به همین دلیل است که چشمان ما گیاهان را به رنگ سبز می‌بیند. شاخص پوشش گیاهی (NDVI) شاخصی است که به منظور توصیف پوشش گیاهی توسعه یافته و با استفاده از تفاوت بین فرسوخ نزدیک (که قویاً توسط گیاهان بازتابش می‌شود) و نور قرمز (که توسط گیاهان جذب می‌شود) می‌تواند نقاط دارای پوشش گیاهی را نمایان سازد. نقشه‌های NDVI به منظور اندازه‌گیری سلامت گیاهان از ترکیب طول‌موج فرسوخ نزدیک و طول‌موج قرمز استفاده می‌کنند [۱۶].

این شاخص مربوط به ظرفیت فتوسنتز و در نتیجه جذب انرژی توسط گیاه است و رابطه آن به صورت رابطه (۱۰) است. به صورت خلاصه شاخص‌های طیفی مرتبط با صخره‌های مرجانی مطابق جدول (۲) آمده است.

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (10)$$

### ۲-۳- دمای درخشندگی<sup>۱</sup>

یکی از ویژگی‌های لندست-۸، داشتن باندهای حرارتی است که با یک ضریب به نام ضریب گسیل<sup>۲</sup> می‌توان به دمای سطح مرتبط شود. پس می‌تواند معیار خوبی برای بررسی دما و بررسی وضعیت صخره‌های مرجانی باشد. برای محاسبه دمای تابشی ابتدا لازم است داده‌های فرسوخ حرارتی به بازتابندگی تبدیل شده و با بکارگیری از عکس قانون پلانک<sup>۳</sup> برای تبدیل بازتابندگی به دمای تابشی لازم است دمای تابشی به دمای واقعی تبدیل شود.

$$SWI = (BLUE * 0.2626) + (GREEN * 0.2141) + (0.0926 * RED) + (NIR * (-0.0656)) + (0.7629 * SWIR1) + (-0.5388 * SWIR2) \quad (8)$$

### ۲-۲-۶- شاخص خودکار استخراج آب (AWEI)

شاخص خودکار استخراج آب (AWEI) جزء شاخص‌هایی است که بیشتر برای استخراج پهنه‌های آبی و وضعیت آب منطقه کاربرد دارد. بنابراین وقتی منطقه مورد نظر ما برای شناسایی سطح آب جزو مناطق شهری و یا جزیره‌ای نباشد این شاخص مورد استفاده قرار نمی‌شود. با اعمال این شاخص در مناطقی که  $AWEI < 0$  پیکسل غیر آب است و در جایی که  $AWEI > 0$  پهنه آبی است. این شاخص در برطرف کردن پیکسل‌های تیره و شناسایی سطوح آبی در مناطق شهری و کوهستانی که مشکل سایه مانع شناسایی درست می‌شود کاربرد دارد. آب در باند ۴ و ۷ لندست-۸ مقدار بازتابندگی کمی دارد و وقتی می‌گوییم پیکسل غیر آب یعنی عوارضی که در این باندها بازتابندگی بالایی دارند و باعث منفی شدن AWEI می‌شوند [۱۵]. سایه در تمام باندها دارای بازتابندگی کم است و رابطه آن مطابق ۹ است.

$$AWEI = GREEN + (2.5 * GREEN) - (1.5 * (NIR - GREEN)) - (0.25 * SWIR) \quad (9)$$

### ۲-۲-۷- شاخص گیاهی تفاضلی نرمال شده (NDVI)

شاخص گرافیکی ساده‌ای است که در تحلیل‌ها و اندازه‌گیری‌های سنجش‌ازدور و ارزیابی وجود یا عدم وجود پوشش گیاهی یک منطقه کاربرد دارد. مطالعه رفتار زمانی پوشش گیاهی، مدل‌سازی آب‌وهوایی، طبقه‌بندی پوشش گیاهی در سطح جهانی، پایش محصولات کشاورزی، مطالعات بیابان‌زدایی و خشک‌سالی، حفاظت محیط زیست، بررسی تعادل میزان انرژی و آب در سطح جهانی از جمله کاربردهای شاخص NDVI است. دامنه تغییرات این شاخص بین +۱ و -۱ است. شیب خط خاک یک بوده و از مبدأ می‌گذرد با افزایش پوشش گیاهی، میزان ارزش حاصل از این نسبت‌گیری نیز افزایش می‌یابد. محدوده این شاخص بین -۱ و +۱ است که مقادیر بالای این شاخص نشان‌دهنده تراکم پوشش گیاهی است. مقادیر مربوط به آب و ابر معمولاً کمتر از صفر است و مقادیر بالاتر نشان‌دهنده پوشش گیاهی هم‌تراکم‌تری است.

<sup>1</sup> Brightness Temperature

<sup>2</sup> Emissivity Coefficient

<sup>3</sup> Plank's law



رابطه (۱۲) مدلی است که غلظت کلروفیل یا رنگ آب را مشخص می‌کند و دارای ضرایب ثابتی است [۱۹].

## ۲-۵- رگرسیون ماشین بردار پشتیبان<sup>۲</sup>

روش SVR در دهه گذشته به‌عنوان یکی از روش‌های قابل اعتماد و کارآمد شناخته شده است. الگوریتم اصلی SVM توسط وپنیک<sup>۳</sup> معرفی شده است. شولکوف<sup>۴</sup> و اسمولا<sup>۵</sup> بیشتر بر روی عمق و نمای کلی از رگرسیون SVM کار کردند. ماشین بردار پشتیبان (SVM) یک روش ماشینی یادگیری به‌منظور دسته‌بندی، رگرسیون و دیگر کارهای آموزشی است. بردارهای پشتیبان به زبان ساده، مجموعه ای از نقاط در فضای n بعدی داده‌ها هستند که مرز دسته‌ها را مشخص می‌کنند و مرزبندی و دسته‌بندی داده‌ها براساس آنها انجام می‌شود و با جابجایی یکی از آن‌ها، خروجی دسته بندی ممکن است تغییر کند. یکی از کاربردهای ماشین بردار پشتیبان، رگرسیون ماشین بردار پشتیبان (SVR) است که به‌طور کلی شامل دو دسته استفاده از داده‌های آموزشی برای به دست آوردن مدل و استفاده از مدل برای پیش‌بینی اطلاعات داده‌های تست است. کیفیت مدل‌های ماشین بردار پشتیبان وابسته به تنظیم مناسب پارامترهای مدل‌سازی ماشین بردار پشتیبان است [۲۰]. در SVR ورودی x در یک فضا با تعداد m عدد ویژگی مختلف به‌صورت غیرخطی است و سپس با استفاده از این ویژگی‌ها یک مدل خطی ساخته می‌شود. مدل خطی به‌صورت رابطه (۱۳) است.

$$f(x, w) = \sum_{i=1}^m w_i g_i(x) + b \quad (13)$$

$g_i(x)$  اشاره به یک مجموعه از انتقال‌های خطی دارد و b ترم بایاس است. کیفیت برآورد به‌وسیله تابع  $L(y, f(x, w))$  اندازه‌گیری شده است. نوع جدیدی از تابع L که تابع  $\varepsilon$  - غیر حساس است را استفاده می‌کند. فرمول این رگرسیون به‌صورت زیر است (رابطه (۱۴))

$$L(y, f(x, w)) = f(x) = \begin{cases} 0 & , |y - f(x, w)| \leq \varepsilon \\ |y - f(x, w)| - \varepsilon & , \text{otherwise} \end{cases} \quad (14)$$

نکته این است که  $\varepsilon$  - غیر حساس منطبق بر کمترین قدر مطلق است. SVR یک رگرسیون خطی در ابعاد بالایی از

جدول (۲) شاخص‌های طیفی.

شاخص	معادل	توصیف
NDWI	Normalized Difference Water Index	شاخص آب تفاضلی نرمال شده
MNDWI	Modification of Normalized Difference Water Index	شاخص آب تفاضلی نرمال شده اصلاحی
GRVI	Green-Red Vegetation Index	شاخص گیاهی سبز-قرمز
FAI	Floating Algae Index	شاخص جلبک شناور
SWI	Surface Wetness Index	شاخص رطوبت سطح
AWEI	Automated Water Extraction Index	شاخص خودکار استخراج آب
NDVI	Normalized Difference Vegetation Index	شاخص گیاهی تفاضلی نرمال شده

از رابطه (۱۱) دمای تابشی محاسبه می‌شود. در این رابطه دمای تابشی برحسب درجه کلون است. R در این بازتابندگی طیفی،  $K_1$  و  $K_2$  ضریب کالیبراسیون بوده و مقدار آن از فایل فراداده لندست-۸ برای دو باند ۱۰ و ۱۱ مشخص است [۱۷].

$$BT = \frac{K2}{\ln\left(\frac{K1}{R} + 1\right)} \quad (11)$$

## ۲-۴- رنگ دریا<sup>۱</sup>

رنگ دریا یک ابزار با ارزش برای نظارت بر کیفیت آب و تغییرات در محیط‌زیست دریایی است، که می‌تواند مستقیماً روی سلامتی انسان خصوصاً در مناطق ساحلی و مرجانی تأثیر بگذارد. مشاهدات رنگ اقیانوس با اندازه‌گیری در طول موج‌های مختلف مرئی محاسبه می‌شود. الگوریتم مورد استفاده در این مقاله الگوریتم OC4 است که با توجه به مطالعات انجام‌شده برای آب‌های نوع دوم یا آب‌های نزدیک ساحل دقت مناسبی دارد [۱۸]. این الگوریتم از یک نسبت انعکاسی با چهار باند بازتابی آبی-سبز استفاده کرده و به صورت رابطه (۱۲) است.

$$C = 10^{\sum_{i=0}^n a_i x^i} \quad (12)$$

$$x = \frac{\log_{10} \{ \text{MAX}\{Rrs(443) \text{ or } Rrs(490) \text{ or } Rrs(510)\} }{Rrs(555)}}{}$$

در این رابطه Rrs بازتابش مربوط به طول موج‌های ۴۴۳ و ۴۹۰ و ۵۱۰ و ۵۵۵ است و در صورت کسر رابطه بیشترین مقدار بازتابش آن‌ها در نظر گرفته می‌شود. C در

<sup>۱</sup> Ocean Colour

<sup>۲</sup> Support Vector Regression

<sup>۳</sup> Vapneik

<sup>۴</sup> Scholkopf

<sup>۵</sup> Smola

## ۲-۶- انتخاب ویژگی<sup>۱</sup>

مساله انتخاب ویژگی، یکی از مسائلی است که در مباحث یادگیری ماشین و شناسایی آماری الگو مطرح است. این مساله در بسیاری از کاربردها (مانند طبقه بندی) اهمیت به سزایی دارد، زیرا در این کاربردها تعداد زیادی ویژگی وجود دارد، که بسیاری از آنها یا بلااستفاده هستند و یا اینکه بار اطلاعاتی چندانی ندارند. حذف نکردن این ویژگی‌ها مشکلی از لحاظ اطلاعاتی ایجاد نمی‌کند ولی بار محاسباتی را برای کاربرد مورد نظر بالا می‌برد. و علاوه بر این باعث می‌شود که اطلاعات غیر مفید زیادی را به همراه داده‌های مفید ذخیره کنیم. در حقیقت انتخاب یک زیرمجموعه از ویژگی‌ها بدون انتقال<sup>۲</sup> و تحول است. این کاهش ابعاد داده‌ها و الگوریتم‌های رگرسیون منجر به اجرای سریع‌تر و مؤثرتر می‌شود. در برخی موارد، ضریب تعیین می‌تواند بهبود یابد؛ در برخی دیگر، منجر به نتیجه واضح‌تر و خلاصه‌تر و تفسیر راحت‌تر از مفهوم مورد نظر شود [۲۱].

به طور کلی روش‌های مختلف انتخاب ویژگی را بر اساس نوع جستجو به دسته‌های مختلفی تقسیم بندی می‌کنند. در بعضی روش‌ها تمام فضای ممکن جستجو می‌گردد. فرآیند انتخاب ویژگی در تمامی روش‌ها را به این بخش‌ها تقسیم می‌کنیم:

۱- تابع تولید کننده<sup>۳</sup>: این تابع زیر مجموعه‌های کاندید را برای روش مورد نظر پیدا می‌کند.

۲- تابع ارزیابی<sup>۴</sup>: زیرمجموعه مورد نظر را بر اساس روش داده شده به آن، ارزیابی و یک عدد به عنوان میزان خوبی روش باز می‌گرداند. روشهای مختلف سعی در یافتن زیرمجموعه‌ای دارند که این مقدار را بهینه کند.

۳- شرط خاتمه: برای تصمیم گیری در مورد زمان توقف الگوریتم.

۴- تابع تعیین اعتبار<sup>۵</sup>: تصمیم می‌گیرد که آیا زیر مجموعه انتخاب شده معتبر است یا خیر؟

الگوریتم‌های متعددی برای این منظور وجود دارد که ما در این مقاله با استفاده از الگوریتم ژنتیک و انتخاب ویژگی ترتیبی جلویی به بررسی ویژگی‌های خود پرداخته‌ایم [۲۲].

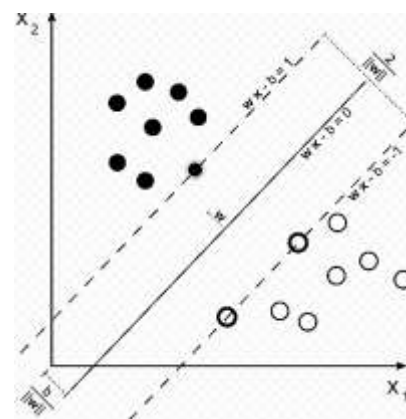
فضای ویژگی‌ها با استفاده از  $\varepsilon$  غیر حساس اجرا می‌کند و سعی می‌کند پیچیدگی مدل را با کمترین  $\|w\|^2$  کاهش دهد. این مسئله می‌تواند به وسیله معرفی متغیر  $\zeta_i^*$  و  $\zeta_i$  و برای  $i=1,2,\dots,n$  اندازه‌گیری انحراف نمونه‌های آموزشی  $\varepsilon$  غیر حساس شرح داده شود. پس SVR به عنوان کمینه کردن تابع به صورت رابطه (۱۵) شود.

$$\min \frac{1}{2} \|w\|^2 + c \sum_{i=1}^n (\zeta_i + \zeta_i^*) \quad (15)$$

$$s. t. = \begin{cases} y_i - f(x_i, w) \leq \varepsilon + \zeta_i^* \\ f(x_i, w) - y_i \leq \varepsilon + \zeta_i \\ \zeta_i, \zeta_i^* \geq 0, i = 1, 2, \dots, n \end{cases} \quad (16)$$

این مسئله بهینه می‌تواند تبدیل به مسئله دوگانه و انتخاب بردارهای حامل و تابع کرنل شود. عملکرد خوب تابع رگرسیون وابسته به انتخاب خوب پارامترهای گاما، اپسیلون ( $\varepsilon$  و  $\gamma$ ) و تابع کرنل است [۱۹]. به دلیل وابسته بودن پارامترهای مدل SVM انتخاب پارامتر بهینه پیچیده است. نرم‌افزارهای اجرایی موجود SVR معمولاً پارامترهای SVM را به عنوان ورودی‌ها با تعریف کاربر اجرا می‌شود. پارامتر  $c$  رابطه جایگزینی بین پیچیدگی مدل و درجه برای این انحرافات بزرگ‌تر از  $\varepsilon$  هموار شود در رابطه (۱۶) بهینه‌سازی S.t را تعیین می‌کند.

پس می‌توان گفت ماشین بردار پشتیبان یا SVM داده‌ها را با توجه به دسته‌های از پیش تعیین شده آنها به یک فضای جدید می‌برد به گونه‌ای که داده‌ها به صورت خطی (یا ابر صفحه) قابل تفکیک و دسته بندی باشند و سپس با یافتن خطوط پشتیبان (صفحات پشتیبان در فضای چند بعدی)، سعی در یافتن معادله خطی دارد که بیشترین فاصله را بین دو دسته ایجاد می‌کند (شکل (۱)). نتایج حاصل از این روش علاوه بر دقت بالا دارای ثبات خوبی هستند [۲۱].



شکل (۱) حل معادله با رگرسیون SVR.

<sup>1</sup> Feature Selection

<sup>2</sup> Transformation

<sup>3</sup> Generation Procedure

<sup>4</sup> Evaluation Function

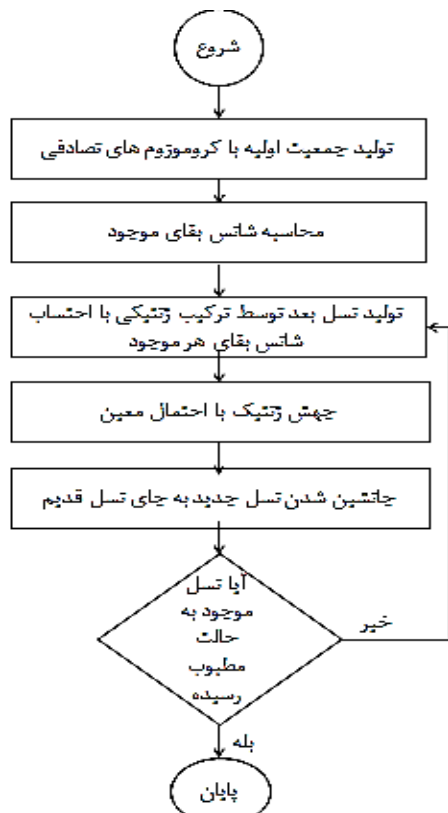
<sup>5</sup> Validation Procedure



۲-۶-۱- الگوریتم ژنتیک (GA)

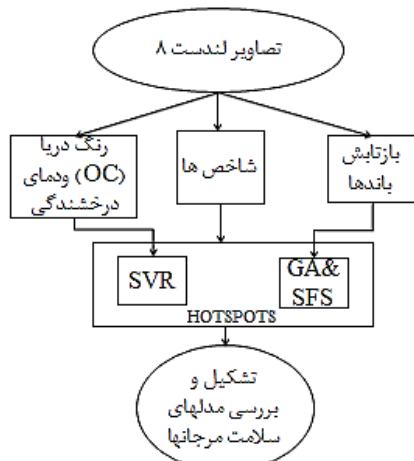
که بهترین عملکرد را داشته باشد به عنوان نمونه انتخاب می‌شود. سپس این عملکرد ویژگی در مقیاس‌ها و سائزهای بزرگ‌تر بررسی می‌شود و این چرخه تا زمانی ادامه می‌یابد که توسعه و پیشرفتی در این ویژگی حاصل نشود و همواره کمترین مقدار خود را داشته باشد. مهم‌ترین نقطه ضعف SFS این است که ویژگی اضافه شده در صورتیکه مناسب نباشد، از مجموعه جواب حذف نمی‌شود [۲۳].

الگوریتم ژنتیک (GA) الگوریتم‌های جستجو با الهام گرفتن از تکامل و انتخاب طبیعی هستند. این الگوریتم را می‌توان برای حل انواع مختلف و متنوع از مشکلات استفاده کرد. الگوریتم ژنتیک به عنوان یک الگوریتم محاسباتی بهینه‌سازی با در نظر گرفتن مجموعه‌ای از نقاط فضای جواب در هر تکرار محاسباتی به نحو مؤثری نواحی مختلف فضای جواب را جستجو می‌کند. این الگوریتم با یک گروه از افراد (کروموزوم) به نام جمعیت شروع می‌شود. هر کروموزوم شامل دنباله‌ای از ژن است که می‌تواند بیت، شخصیت‌ها، و یا اعداد باشد [۲۲].



شکل (۲) فلوجارت الگوریتم ژنتیک.

الگوریتم‌های ژنتیک، جستجوی اصلی را در فضای پاسخ به اجرا می‌گذارند. این الگوریتم‌ها با تولید نسل آغاز می‌شوند که وظیفه ایجاد مجموعه نقاط جستجوی اولیه به نام جمعیت اولیه را بر عهده دارند. از آنجایی که الگوریتم‌های ژنتیک برای هدایت عملیات جستجو به طرف نقطه بهینه از روش‌های آماری استفاده می‌کنند، در فرآیندی که به انتخاب طبیعی وابسته است، جمعیت موجود به تناسب برانزندی افراد آن نسل بعد انتخاب می‌شود. سپس عملگرهای ژنتیکی شامل: انتخاب، پیوند (ترکیب)، جهش و دیگر عملگرهای احتمالی اعمال شده و جمعیت جدید به وجود می‌آید. پس از آن جمعیت جدیدی جایگزین جمعیت پیشین می‌شود و این چرخه ادامه می‌یابد. معمولاً جمعیت جدید برانزندی بیشتری دارد این بدان معناست که از نسلی به نسل دیگر جمعیت بهبود می‌آید. هنگامی جستجو نتیجه بخش خواهد بود که به حداکثر نسل ممکن رسیده باشیم یا همگرایی حاصل شده باشد و یا معیارهای توقف برآورده شده باشد (شکل (۲)) [۲۲].



شکل (۳) فلوجارت روند اجرای مقاله.

۲-۶-۲- الگوریتم انتخاب ویژگی ترتیبی<sup>۱</sup>

به طور کلی اغلب این الگوریتم شامل دو بخش انتخاب جلویی ترتیبی<sup>۲</sup> (SFS) و انتخاب عقبی ترتیبی<sup>۳</sup> (SBS) است. در SFS که ساده‌ترین الگوریتم جستجوگر است با نسبت دادن تهی به تمامی مجموعه ویژگی‌ها کل ویژگی‌ها به وسیله یک ویژگی ارزیابی می‌شود. سپس در نزدیکی این ویژگی

<sup>1</sup> Sequential Feature Selection

<sup>2</sup> Sequential Forward Selection

<sup>3</sup> Sequential Backward Selection

### ۳-۲- داده‌های میدانی و ماهواره‌ای مورد استفاده

داده‌های مورد استفاده شامل داده‌های دمای سطح آب به همراه موقعیت جغرافیایی صخره‌های مرجانی برای منطقه مورد مطالعه (۶۰ نقطه برای آموزش و ۳۰ نقطه برای تست) از سایت‌های مربوطه اخذ گردید [۲۶]. همچنین محصولات ماهانه مدیس (دمای سطح آب) برای بدست آوردن بیشترین متوسط ماهیانه دمای سطح آب (۲۴,۹۰) درجه سانتیگراد) مربوط به الگوریتم‌های نظارت بر صخره‌های مرجانی و تصاویر لندست-۸ برای نقاط مدنظر است. با بررسی و حذف داده‌های زائد در نهایت با ۶۵٪ داده‌ها (۶۰ نقطه) برای آموزش مدل و مابقی (۳۰ نقطه) به تست مدل با استفاده از SVR پرداختیم. همان‌طور که قبلاً گفته شد برای تشکیل مدل نیاز به تعیین پارامترهای کرنل گوسی شامل C و  $\gamma$  و  $\epsilon$  برای مدل خود هستیم که به‌طور کلی به دو صورت انجام می‌گیرد. اعتبار مقاطع<sup>۲</sup> (CV) و جستجو شبکه<sup>۳</sup> (GS) [۲۷].

ما در این پروژه با استفاده از روش جستجو شبکه به تعیین پارامترهای خود برای مدل‌سازی پرداختیم. الگوریتم جستجوی شبکه‌ای، روش رایج در انتخاب پارامترهای بهینه مدل است.

در این روش، شبکه‌ای k بعدی بر روی محدوده پارامترها قرار می‌گیرد که k تعداد پارامترهای مجهول می‌باشد، سپس کیفیت تمام مجموعه جواب‌های ممکن در نقاط شبکه ارزیابی می‌شود و آن دسته از پارامترهایی که کمترین خطای طبقه‌بندی را دارند، به‌عنوان پارامترهای ماشین‌های بردار پشتیبان انتخاب می‌شوند. به علت پیوسته بودن مقادیر پارامترهای موردنظر، برای رسیدن به دقت بالا باید شبکه‌ای با تراکم بالا در نظر گرفته شود که بررسی تمام این نقاط شبکه زمان محاسبات را به شدت افزایش می‌دهد [۱۹ و ۲۲].

با توجه به محدودیت‌های روش جستجوی شبکه‌ای در زمان و محاسبات، الگوریتم‌های بهینه‌سازی مختلفی برای حل این مسئله در نظر گرفته شده‌اند: الگوریتم‌های خرد جمعی، شبیه‌سازی تبرید تدریجی و الگوریتم ژنتیک. در ادامه روند ریاضی انتخاب پارامتر c شرح داده شده است.

$$|f(x)| = \left| \sum_{i=1}^{n_{sv}} (a_i - a_i^*) k(x_i, x) \right| \quad (17)$$

در عوض در SBS این روش برعکس SFS کارش را با مجموعه‌ای شامل تمام ویژگی‌ها شروع می‌کند و در هر بار تکرار الگوریتم، ویژگی که بوسیله تابع ارزیابی انتخاب می‌شود، را از مجموعه مورد نظر حذف می‌کند. این کار را تا زمانی ادامه می‌دهد که تعداد ویژگی‌ها برابر یک تعداد معینی شود. مانند روش قبل مشکل این روش اینست که ویژگی حذف شده را دیگر به مجموعه اضافه نمی‌کند، حتی اگر مناسب باشد [۲۴]. با توجه به روش‌ها و الگوریتم‌های صحبت شده روند کلی و فلوچارت کلی این مقاله مطابق شکل (۳) است.

### ۳- پیاده‌سازی و نتایج

#### ۳-۱- معرفی منطقه مورد مطالعه

جزیره مرجانی هرون<sup>۱</sup> به طول ۸۰۰ متر و بیشترین عرض ۳۰۰ متر و به مساحت ۱۶ هکتار در مختصات جغرافیایی ۱۵۱,۹۰۸۳۹۷ و ۲۳,۴۳۸۵۳۱- دارای مجموعه عظیمی از صخره‌های مرجانی است که در شرق کشور استرالیا و در اقیانوس آرام در نزدیکی مدار رأس الجدی قرار دارد. صخره‌های مرجانی این جزیره بخشی از پارک دریایی GBR در استرالیا است که سالانه دو میلیارد توریست را با خود به این مکان می‌آورد و قطعاً پربازدیدترین جاذبه گردشگری در این کشور است. جاذبه‌ای طبیعی و بی‌نظیر که تا فاصله‌ای طولانی در ساحل استرالیا از بوندابریگ تا رأس شمالی کوئینزلند امتداد یافته است.

دیواره بزرگ مرجانی، بزرگ‌ترین ساحل مرجانی دنیا، زیستگاه بیش از ۱۵۰۰ نوع ماهی، ۴۱۱ نوع مرجان سخت، همچنین یک‌سوم مرجان‌های نرم جهان و ۱۳۴ گونه کوسه است؛ این صخره‌های مرجانی در سال ۱۹۸۱ جزو میراث جهانی یونسکو ثبت شدند و مقصد گردشگری-تفریحی بی‌مانندی برای ماجراجویان و علاقه‌مندان به ورزش‌های آبی فراهم نموده است. در سال ۲۰۱۲ اعلام شد که تعداد مرجان‌های بخشی از دیواره بزرگ مرجانی بزرگ طی ۲۷ سال به نصف کاهش یافته است [۲۵]. به گفته دانشمندان دلیل کاهش تعداد این صخره‌های دریایی، توفان‌های بسیار و افزایش دمای سطح زمین و آب دریاها و افزایش جمعیت ستاره‌های دریایی است که غذای این مرجان‌ها را می‌مکد [۲۵].

<sup>2</sup> Cross Validation

<sup>3</sup> Grid-Search

<sup>1</sup> Heron Island

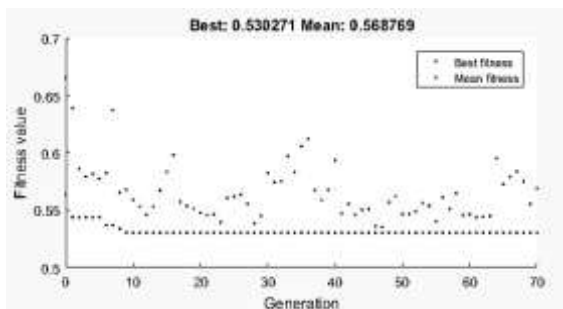
(شکل ۵)) دارای ۱۷ ویژگی برای مسئله مربوطه هستیم که شماره‌های ۱ و ۲ دمای درخشندگی باندهای ۱۰ و ۱۱، شماره‌های ۳ تا ۶ شاخص‌های NDVI, NDWI, GRVI, MNDWI هستند. شماره‌های ۷ تا ۱۳ مربوط به بازتابش باندهای ۱ تا ۷ لندست-۸ است. شماره ۱۴ مربوط به کمیت رنگ دریا، ۱۵ مربوط به FAI، ۱۶ مربوط به AEWI و ۱۷ مربوط به SWI است. این نکته حائز اهمیت است که ویژگی‌های شماره‌های ۱ و ۳ و ۷ و ۱۱ و ۱۲ و ۱۵ بهترین ویژگی‌ها و شاخص‌ها برای بررسی وضعیت سلامت مرجان‌ها برای منطقه مورد نظر هستند.

درواقع سلامت مرجان‌ها تابعی مطلوب‌تر از ویژگی‌های دمای درخشندگی باند ۱۰، شاخص NDVI، بازتابش باند ۱ (باند ساحلی)، بازتابش باند ۵ (فروسرخ نزدیک)، باند ۶ (فروسرخ کوتاه) و شاخص FAI هستند (رابطه ۲۲)).

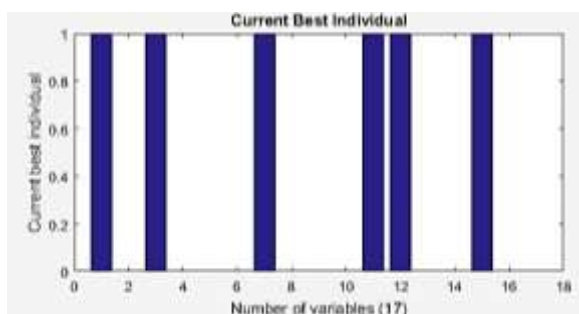
$$CR = f(BT10, NDVI, RAD1, RAD5, RAD6, FAI) \quad (22)$$

در مورد الگوریتم SFS می‌توان خاطرنشان کرد این الگوریتم ویژگی بازتابش باند ۵ (فروسرخ نزدیک) و شاخص جلبک‌های شناور را مناسب برای ارزیابی و بررسی مدل وضعیت سلامت مرجان‌ها نشان داده است (رابطه ۲۳).

$$CR = f(RAD5, FAI) \quad (23)$$



شکل (۴) همگرا شدن روش SVR-GA.



شکل (۵) بررسی ویژگی‌های انتخاب شده در الگوریتم ژنتیک.

$$\leq \sum_{i=1}^{n_{sv}} |a_i - a_i^*| \cdot |k(x_i, x)| \quad (18)$$

$$\leq \sum_{i=1}^{n_{sv}} c_i \cdot |k(x_i, x)| \quad (19)$$

در رابطه (۱۷) تعداد بردارهای حامل است و  $k$  همان تابع کرنل است و ادامه روابط به صورت (۱۸) و (۱۹) است. رابطه (۲۰) در روابط بالا جایگذاری می‌شود.

$$k(x_i, x) = \exp\left(-\frac{\|x-x_i\|^2}{2p^2}\right) \quad (20)$$

با استفاده از فرمول (۲۱) مقدار  $C$  محاسبه می‌شود.

$$C = \max(|\bar{y} + 3\sigma_y|, |\bar{y} - 3\sigma_y|) \quad (21)$$

در رابطه (۲۱) میانگین پاسخ‌هایی آموزشی و  $\sigma_y$  انحراف استاندارد پاسخ‌های آموزشی است [۱۹ و ۲۷]. به منظور انتخاب مقادیر  $\gamma$  و  $\epsilon$  برای هر پارامتر محدوده  $2^n$  به صورتی که  $n = 7, 6, \dots, 1$  انتخاب شد و مورد بررسی قرار گرفت که با استفاده از این روش و بررسی بازه‌های مختلف  $\gamma, \epsilon, C$  به ترتیب ۸، ۰، ۱۲۵، ۲، ۴۴۶، بدست آمد. سپس از الگوریتم GA و الگوریتم انتخاب ترکیبی جلویی (SFS) در ترکیب با SVR برای تعیین بهترین ویژگی‌های به دست آمده از تصاویر لندست-۸ برای بازیابی وضعیت سلامت مرجان‌ها استفاده شد.

### ۳-۳- نتایج

همان‌طور که گفته شد با استفاده از الگوریتم ژنتیک به بررسی ویژگی‌های اشاره شده پرداختیم و در نهایت ترکیب الگوریتم ژنتیک استفاده شده و SVR به همگرایی رسید (شکل ۴)). پارامترهای الگوریتم ژنتیک استفاده شده شامل جمعیت اولیه<sup>۱</sup> و تعداد دفعات تکرار الگوریتم<sup>۲</sup> (تعداد نسل‌ها) و نرخ جهش<sup>۳</sup> و نسبت عملگر تقاطع<sup>۴</sup> الگوریتم به ترتیب ۲۰، ۷۰٪، ۵٪ و ۷۰٪ هستند و معیار بهینه کردن ویژگی‌ها کمینه ساختن RMSE است. در روش ژنتیک هر عضو جمعیت به وسیله یک رشته از ۰ و ۱ به طول تعداد باندهای تصویر و ویژگی‌ها و شاخص‌ها نمایش داده می‌شود. در این کروموزوم، مقدار بیت ۰ به معنای حذف ویژگی متناظر و ۱ به معنای انتخاب آن می‌باشد.

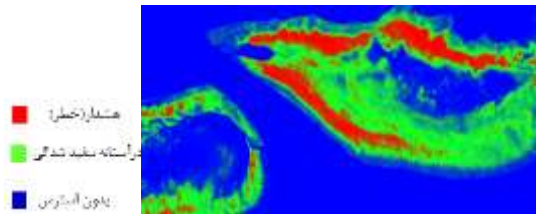
این الگوریتم مناسبترین رشته‌ها را از میان اطلاعات تصادفی سازماندهی شده انتخاب می‌کند. مطابق خروجی الگوریتم ژنتیک، که ویژگی‌های ما در آن نشان داده شده

<sup>1</sup> Population No

<sup>2</sup> Max Iterations

<sup>3</sup> Mutation Ratio

<sup>4</sup> Crossover Ratio



شکل (۷) طبقه‌بندی وضعیت سلامت صخره‌های مرجانی براساس مدل پیشنهادی SVR-SFS.

#### ۴- نتیجه‌گیری

در نهایت می‌توان گفت در این مطالعه یک روش برای بررسی وضعیت سلامت مرجان‌ها با استفاده از ویژگی‌ها و شاخص‌های مرتبط با آن ارائه کرده است. نتایج نشان می‌دهد که این الگوریتم قادر به دست آوردن اطلاعات مورد نظر با دقت و کارایی بالا ( $R^2 = 0.979$  و  $RMSE = 0.591$ ) است. این روش قادر به استفاده از مدلی ترکیبی از ویژگی‌های مختلف با استفاده از الگوریتم ژنتیک و الگوریتم انتخاب ترتیبی جلویی و SVR و سپس حل مشکل در فضا با ابعاد بالاتر است که در حالت‌ها مدل پیشنهادی SVR-GA به بهترین حالت برای مدل خود رسیدیم ( $R^2 = 0.983$  و  $RMSE = 0.53$ ) پس با ترکیبی از الگوریتم‌ها و رگرسیون پشتیبان بردار و تنظیم پارامترهای مرتبط با این روش با استفاده از جستجوی شبکه‌ای، به یک ترکیب با ویژگی‌های مطلوب برای بررسی سلامت مرجان‌ها رسیدیم. ضمن اینکه با استفاده از الگوریتم ژنتیک نشان دادیم ویژگی‌های دمای درخشندگی باند ۱۰، شاخص NDVI، بازتابش باند ۱ (باند ساحلی)، بازتابش باند ۵ (فروسرخ نزدیک)، باند ۶ (فروسرخ کوتاه) و شاخص FAI بهترین‌ها برای بررسی وضعیت سلامت مرجان‌ها برای منطقه مورد مطالعه هستند. در این مقاله مجموعه داده‌های مورد مطالعه در مقیاس منطقه‌ای است و به منظور پژوهش و مطالعه در این زمینه پیشنهاد می‌شود که پیاده‌سازی این روش در مقیاس جهانی برای رسیدن به الگویی جهانی همراه با دقت بالا در دستور کار افراد قرار گیرد.

#### ۵- مراجع

[1] امیر مسعود چگونیان، "بهبود دقت روش‌های طبقه‌بندی اکوسیستم‌های مرجانی در تصاویر ماهواره‌ای خلیج فارس"، پایان‌نامه کارشناسی ارشد در رشته مهندسی عمران-

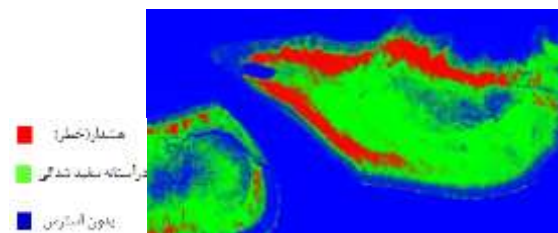
جدول (۴) تمامی معیارهای دقت داده‌های آزمایشی و تست را نشان می‌دهد. همان‌طور که از جدول (۳) معلوم است خطای مجذور میانگین ریشه‌ای برای داده‌های آزمایش برابر ۰.۵۹۱ و برای داده‌های آموزشی ۰.۱۲۳ شده است که بیانگر دقت مناسبی است. در جدول (۳) MSE خطای میانگین ریشه‌ای، NRMSE خطای مجذور میانگین ریشه‌ای نرمال شده و  $R^2$  ضریب تعیین هستند.

می‌توان وضعیت سلامت مرجان‌ها را با توجه به مدل‌سازی که توسط SVR انجام گرفته و به کمک الگوریتم ژنتیک که بهترین ویژگی‌های آن استخراج شده را به صورت زیر در ۳ کلاس طبقه بندی کرد (شکل (۶)). برای منطقه مورد نظر می‌توان فهمید چه مناطقی کلاسی شامل (۱) هشدار و خطر، (۲) در آستانه سفیدشدگی و مردن مرجانها و (۳) مناطق بدون استرس، دارند که در این حالت به  $R^2 = 0.983$  و  $RMSE = 0.53$  رسیدیم که به اندازه ۰.۰۰۴ بهبود را در مدل‌سازی شاهد هستیم (نسبت به حالت SVR).

جدول (۳) معیارهای دقت برای داده‌های آزمایشی و تست.

روش	MSE	RMSE	NRMSE	$R^2$
آموزش	۰.۰۱۵	۰.۱۲۳	۰.۸۴۱	۰.۷۱۶
تست	۰.۳۴۹	۰.۵۹۱	۰.۱۵	۰.۹۷۹

همچنین وضعیت سلامت مرجانها برای منطقه مورد مطالعه با توجه به مدل‌سازی که توسط SVR و به کمک الگوریتم انتخاب ویژگی ترتیبی جلویی (به منظور استخراج بهترین ویژگی‌ها) انجام شد را به صورت زیر در ۳ کلاس طبقه بندی کردیم. (شکل (۷)) که برای اکثر مناطق دارای کلاسی یکسان با روش پیشنهادی قبلی دارد در این حالت به  $R^2 = 0.98$  و  $RMSE = 0.47$  رسیدیم که اندازه ۰.۰۰۱ بهبود در مدل ساخته شده نسبت به حالت کلی SVR را داریم.



شکل (۶) طبقه‌بندی وضعیت سلامت صخره‌های مرجانی براساس مدل پیشنهادی SVR-GA.

- [13] Hu, Chuanmin. "A Novel Ocean Color Index to Detect Floating Algae in the Global Oceans", *Remote Sensing of Environment*, Vol.113, No.10, pp.2118-2129, 2009.
- [14] Motohka, Takeshi, et al. "Applicability of Green-red Vegetation Index for Remote Sensing of Vegetation Phenology", *Remote Sensing*, Vol.2, No.10, pp.2369-2387, 2010.
- [15] Campos, João C., Neftalí Sillero, and José C. Brito. "Normalized Difference Water Indexes Have Dissimilar Performances in Detecting Seasonal and Permanent Water in the Sahara-Sahel Transition Zone", *Journal of Hydrology*, Vol.464, pp.438-446, 2012.
- [16] Rokni, K., et al. "Water Feature Extraction and Change Detection Using Multitemporal Landsat Imagery. *Remote Sensing*", Vol.6, pp.4173-4189, 2014.
- [17] Kogan, F. N., "Application of Vegetation Index and Brightness Temperature for Drought Detection", *Advances in Space Research*, Vol.15, No.11, pp.91-100, 1995.
- [18] E. T. Harvey, S. Kratzer, and P. Philipson, "Satellite-based Water Quality Monitoring for Improved Spatial and Temporal Retrieval of Chlorophyll-a in Coastal Waters," *Remote Sensing of Environment*, Vol.158, pp.417-430, 2015.
- [19] نسرین مرادی "بررسی و مدل‌سازی رنگ اقیانوس با استفاده از تصاویر ماهواره‌ای با توان تفکیک مکانی بالا". پایان‌نامه کارشناسی ارشد در رشته مهندسی عمران- نقشه‌برداری گرایش هیدروگرافی\_دانشگاه تهران (۱۳۹۵).
- [20] Üstün, B., et al. "Determination of Optimal Support Vector Regression Parameters by Genetic Algorithms and Simplex Optimization", *Analytica Chimica Acta*, Vol.544, No.1, pp.292-305, 2005.
- [21] Nikraftara, Z., M. Hasanloua, and M. Esmaeilzadeha. "Novel Snow Depth Retrieval Method Using Time Series Ssmi Passive Microwave Imagery", *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pp.525-530, 2016.
- [22] Zahir, N., and H. Mahdi. "Snow Depth Estimation Using Time Series Passive Microwave Imagery via Genetically Support Vector Regression (case Study Urmia Lake Basin)", *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol.40, No.1, p.555, 2015.
- [23] Aha, David W., and Richard L. Bankert. "A Comparative Evaluation of Sequential Feature Selection Algorithms", *Learning from Data*. Springer New York, pp.199-206, 1996.
- [24] Jain, Anil, and Douglas Zongker. "Feature Selection: Evaluation, Application, and Small Sample Performance", *IEEE transactions on*
- نقشه‌برداری گرایش سنجش از دور، دانشگاه خواجه نصیر  
طوسی (۱۳۹۳)
- [2] Sanghvi, D., "Inventory and Monitoring of Coral Reefs of United Arab Emirates (UAE), Arabian Gulf, Using Remote Sensing Techniques Lauer, Matthew, and Shankar Aswani, Integrating Indigenous Ecological Knowledge and Multi-Spectral Image Classification for Marine Habitat Mapping in Oceania", *Ocean & Coastal Management*, Vol. 51, No.6, pp.495-504, 2008.
- [3] Smith, V. Elliott, Robert H. Rogers, and Larry E. Reed, "Automated Mapping and Inventory of Great Barrier Reef zonation with LANDSAT data", *OCEAN 75 Conference. IEEE*, 1975.
- [4] Peddle, Derek R., Ellsworth F. LeDrew, and Heather M. Holden. "Spectral Mixture Analysis of Coral Reef Abundance from Satellite Imagery and in Situ Ocean Spectra, Savusavu Bay, Fiji." *Proceedings of the Third Thematic Conference on Remote Sensing for Marine and Coastal Environments*. Vol.2. 1995.
- [5] Hochberg, Eric J., and M. J. Atkinson. "Spectral Discrimination of Coral Reef Benthic Communities", *Coral reefs*, Vol.19, No.2, pp.164-171, 2000.
- [6] Andréfouët, Serge, et al. "Multi-site Evaluation of IKONOS Data for Classification of Tropical Coral Reef Environments", *Remote sensing of environment*, Vol.88, No.1, pp.128-143, 2003.
- [7] G. Camps-Valls and L. Bruzzone, "Kernel-based Methods for Hyperspectral Image Classification", *IEEE Transactions on Geoscience and Remote Sensing*, Vol.43, pp.1351-1362, 2005.
- [8] Roberts, J. Murray, "Cold-water Corals: the Biology and Geology of Deep-Sea Coral Habitats". Cambridge University Press, 2009.
- [9] Knudby, Anders, et al. "Mapping Coral Reef Resilience Indicators using Field and Remotely Sensed Data" *Remote Sensing*, Vol.5, No.3, pp.1311-1334, 2013.
- [10] Liu, Gang, et al. "Reef-Scale Thermal Stress Monitoring of Coral Ecosystems: New 5-km Global Products from NOAA Coral Reef Watch" *Remote Sensing*, Vol.6, No.11, pp.11579-11606, 2014.
- [11] Zahir, N., and H. Mahdi. "Snow Depth Estimation Using Time Series Passive Microwave Imagery via Genetically Support Vector Regression (case Study Urmia Lake Basin)." *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol.40, No.1, p.555, 2015.
- [12] Feyisa, Gudina L., et al. "Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery" *Remote Sensing of Environment*, Vol.140, pp.23-35, 2014.

- Pattern Analysis and Machine intelligence,  
Vol.19, No.2, pp.153-158, 1997.
- [25] Authority, Great Barrier Reef Marine Park.  
"Great barrier reef outlook report 2014."
- [26] <https://www.coris.noaa.gov/>
- [27] Bao, Yukun, and Zhitao Liu. "A Fast Grid  
Search Method in Support Vector Regression  
Forecasting Time Series", International  
Conference on Intelligent Data Engineering and  
Automated Learning, Springer Berlin  
Heidelberg, 2006.