

بررسی تاثیر و نقش عوامل توپوگرافی و آب و هوایی در تهیه نقشه آسیب پذیری بیماری لپتوسپیروز

مهرداد آهنگرگانی^۱، مهدی فرنقی^{۲*}، محمدرضا شیرزادی^۳

- ۱- دانشجوی کارشناسی ارشد سیستم‌های اطلاعات مکانی دانشکده مهندسی ژئودزی و ژئوماتیک، دانشگاه صنعتی خواجه‌نصیرالدین طوسی
- ۲- استادیار گروه سیستم‌های اطلاعات مکانی دانشکده مهندسی ژئودزی و ژئوماتیک، دانشگاه صنعتی خواجه‌نصیرالدین طوسی
- ۳- رئیس مرکز کنترل و پیشگیری بیماری‌های واگیر وزارت بهداشت، درمان و آموزش پزشکی

تاریخ دریافت مقاله ۱۳۹۵/۰۸/۰۵ تاریخ پذیرش مقاله: ۱۳۹۶/۰۳/۱۰

چکیده

بیماری لپتوسپیروز یکی از رایج‌ترین و گسترش یافته‌ترین بیماری عفونی مشترک میان انسان و حیوان در سراسر دنیا بشمار می‌رود که به وسیله باکتری بیماری‌زایی به نام لپتوسپیرا ایجاد می‌شود و به شکل مستقیم یا غیر مستقیم از حیوان به انسان منتقل می‌شود. این بیماری در تمام مناطق جهان اتفاق می‌افتد، اما در مناطق گرمسیر پر باران شایع‌تر است و در اغلب این مناطق بومی می‌باشد. هدف اصلی این تحقیق، بررسی فضلی ارتباط میان متغیرهای محیطی و توپوگرافی از قبیل ارتفاع، شیب، پوشش گیاهی، متوسط دمای هوا، متوسط رطوبت هوا، میزان بارندگی و تعداد روزهای یخبندان با وقوع بیماری لپتوسپیروز و تهیه نقشه آسیب‌پذیری بیماری لپتوسپیروز در سطح بخش‌های استان‌های شمالی ایران می‌باشد. در این تحقیق، تحلیل همبستگی پیرسون به منظور بررسی نوع و شدت همبستگی بین متغیرهای محیطی و توپوگرافی با بروز بیماری لپتوسپیروز و توصیف ارتباط بین بروز فضلی بیماری لپتوسپیروز و متغیرهای محیطی و توپوگرافی مورد استفاده قرار گرفت. نتایج حاصل از تحلیل همبستگی پیرسون مشخص نمود که متغیرهای میانگین رطوبت هوا، میانگین دمای هوا و میزان بارندگی مهم‌ترین فاکتورها بر شیوع بیماری لپتوسپیروز در منطقه می‌باشند. همچنین نتایج حاصل از بررسی‌های آماری نشان داد که بیش‌ترین موارد بیماری در فصل تابستان و اواخر فصل بهار به ثبت رسیده و کم‌ترین رخداد بیماری در فصل زمستان رخ داده است. علاوه بر این، نتایج آماری به‌دست آمده حاکی از وجود موارد وقوع بیماری لپتوسپیروز در نواحی مرکزی استان گیلان، مناطق شرقی استان مازندران و مناطق غربی استان گلستان، با آب و هوای معتدل و مرطوب با بارش شدید باران می‌باشد. در نهایت روش ماشین بردار پشتیبان به منظور مدل‌سازی ارتباط بین متغیرهای محیطی و توپوگرافی و توزیع فضلی الگوی مکانی بیماری لپتوسپیروز مورد استفاده قرار گرفت و عملکرد این مدل با استفاده از معیار ارزیابی منحنی مشخصه عملکرد سیستم مورد بررسی قرار گرفت. نتایج حاصل از ارزیابی نقشه‌های خروجی روش ماشین بردار پشتیبان نشان دادند که این روش عملکرد بسیار عالی در تهیه نقشه آسیب‌پذیری بیماری لپتوسپیروز داشته است.

واژه‌های کلیدی: تحلیل پیرسون، سامانه اطلاعات مکانی، ماشین بردار پشتیبان، منحنی مشخصه عملکرد سیستم، بیماری لپتوسپیروز

* نویسنده مکاتبه کننده: تهران - خیابان ولیعصر - تقاطع میرداماد غربی - دانشکده ژئودزی و ژئوماتیک - دانشگاه صنعتی خواجه نصیرالدین طوسی.

تلفن: ۰۹۱۲۲۴۳۲۸۴۷

۱- مقدمه

بیماری لپتوسپیروز به دلیل اپیدمی و شیوع فزاینده، مشکلی جهانی و مسئله‌ای برای سلامت عمومی در کشورهای در حال توسعه و توسعه یافته می‌باشد. در مناطق پر باران حاره و تحت حاره که اغلب کشورهای در حال توسعه در آنجا وجود دارند، نسبت به مناطق معتدل، این مشکل به صورت جدی‌تری خودنمایی می‌کند [۱]. در واقع این بیماری در تمام مناطق جهان اتفاق می‌افتد اما در مناطق گرمسیر پر باران شایع‌تر است و در اغلب این مناطق بومی^۱ می‌باشد [۲]. اگرچه اهمیت و بزرگی مسئله در مناطق حاره و تحت حاره به آب و هوا و شرایط محیطی نسبت داده می‌شود اما به احتمال زیاد برخورد با محیط آلوده به لپتوسپیروز که در اثر روش‌های کشاورزی محلی، سکونت نامناسب و دفع غیربهداشتی زباله‌ها و مواد زاید که همگی باعث ایجاد منابع متعدد عفونت و آلودگی می‌شوند را نباید از یاد برد. در کشورهای دارای آب و هوای معتدل علاوه بر وجود بیماران با انتقال در محل، ممکن است از طریق مسافرانی که از مناطق حاره به این مناطق می‌آیند نیز بر تعداد بیماران افزوده شود [۳]. این بیماری یکی از رایج‌ترین و مخوف‌ترین بیماری عفونی مشترک میان انسان و حیوان در سراسر دنیا بشمار می‌رود که به وسیله باکتری بیماری‌زایی به نام لپتوسپیرا ایجاد می‌شود و به شکل مستقیم یا غیر مستقیم از حیوان به انسان منتقل می‌شود. منبع اصلی این بیماری جوندگان و حیوانات وحشی می‌باشند که باکتری لپتوسپیرا را در فضولات خود دفع می‌کنند. برخی از حیوانات اهلی همچون سگ و گاو نیز در انتشار عفونت این بیماری موثرند و به عنوان مخزن عفونت برای انسان و دام‌ها عمل می‌کنند [۴]. در سال‌های اخیر، سیر تحولی بیماری لپتوسپیروز از

یک بیماری سنتی روستایی که باعث افزایش شیوع بیماری در میان کشاورزان، شالیکاران، ماهیگیران و غیره می‌شد، تغییر یافته و به عنوان عامل اپیدمی‌ها در جوامع شهری با سطح بهداشت نامناسب و ضعیف، تبدیل شده است [۵]. بیماری لپتوسپیروز در کشورهای در حال توسعه مثل ایران یک بیماری شغلی محسوب می‌شود و بیشتر در میان کشاورزان، دامداران، کارکنان کشتارگاه‌ها، قصاب‌ها و ماهیگیران در فصول گرم سال رخ می‌دهد [۶]. بیماری لپتوسپیروز اغلب گسترش فصلی دارد و با افزایش بارندگی و افزایش درجه حرارت بر شیوعش افزوده می‌شود اما بیماری در تمام طول سال می‌تواند اتفاق بیافتد. هرکجا که انسان در تماس با ادرار و فضولات حیوان، آب، خاک و محیط آلوده قرار بگیرد، احتمال وقوع این بیماری وجود دارد. مخصوصاً در مناطقی از جهان که شرایط آب و خاک آن قلیایی باشد. به همین علت می‌باشد که در سال‌های اخیر در استان‌های شمالی ایران به دلیل دارا بودن آب و هوای معتدل، مرطوب و آب و خاک قلیایی موارد قابل توجهی از بیماری انسانی مشاهده شده است [۳ و ۷].

از آنجائی که عامل مکان به‌عنوان یک عامل مهم و کلیدی در مطالعات مربوط به بهداشت عمومی و همه‌گیری بیماری‌ها به‌شمار می‌آید [۸، ۹] و همچنین سامانه‌های اطلاعات مکانی (GIS^۲) قابلیت‌های زیادی در مورد مدلسازی پدیده‌های مکانی دارند [۱۰]. لذا چنین سامانه‌هایی می‌توانند در زمینه مطالعه و بررسی بیماری‌ها و عوامل موثر بر آن‌ها بسیار مفید واقع شوند و با استفاده از اطلاعات جمع‌آوری شده به تجزیه و تحلیل مکانی در مورد آن‌ها بپردازند [۱۱]. با توجه به موارد اشاره شده، لزوم بررسی مکانی در زمینه بیماری لپتوسپیروز، مخازن و ناقلین آنها به‌منظور در دست داشتن آخرین اطلاعات لازم و به‌کار بردن آنها در برنامه‌های کنترلی این بیماری،

^۲ Geographical Information System^۱ Endemic

ایدز و بسیاری دیگر از ناشناخته‌های موثر در ایجاد بیماری‌های مختلف، از جمله دستاوردهای پژوهش‌های مکانی مرتبط با بیماری در سراسر جهان و از نمونه‌های برجسته تاریخی این نوع پژوهش‌های همه‌گیرشناسی هستند. امروزه سامانه اطلاعات مکانی به عنوان بخشی از فناوری اطلاعات توسط سازمان‌های مدیریت بهداشت و درمان، در سطوح مختلف جهانی، ملی، منطقه‌ای و محلی به طور گسترده مورد استفاده قرار گرفته است [۱۳]. در سطح بین‌المللی با استفاده از سامانه اطلاعات مکانی و تحلیل‌های مکان-آمار متعدد، مطالعات گسترده‌ای از سراسر جهان در مورد بیماری‌های واگیر مشترک میان انسان و حیوان گزارش شده است که به بررسی ارتباط بین پارامترهای محیطی و آب و هوایی با شیوع چنین بیماری‌ها پرداخته‌اند. به‌عنوان نمونه، بارسلوز و سابروزا در سال ۲۰۰۱ میلادی، با استفاده از ترکیب داده‌های محیطی و به‌کارگیری تحلیل‌های مکانی، تاثیر عوامل محیطی بر بیماری لپتوسپیروز را در کشور برزیل در تابستان ۱۹۹۶ مورد بررسی قرار دادند. نتایج حاصل از مطالعه نشان دادند که بیشترین نرخ بیماری درون مناطق آسیب‌پذیر نسبت به سیل و در اطراف محل‌های جمع‌آوری زباله رخ می‌دهند [۱۴]. مطالعه انجام شده در برزیل توسط باویا و همکاران با استفاده از سامانه اطلاعات مکانی و سنجش از دور (RS^۱) در مورد بیماری لیشمانیوز احشایی نشان داده است که شاخص پوشش گیاهی نرمال شده (NDVI^۲) یکی از مهمترین عوامل محیطی موثر بر شیوع بیماری لیشمانیوز احشایی بوده است. بدین‌صورت که بیش‌تر ناقلین در مطالعه فوق در مناطق دارای پوشش گیاهی کم مشاهده شده‌اند [۱۵]. همچنین پزشکی و همکاران به بررسی فاکتورهای بیماری وبا از سال ۱۹۹۷ تا ۲۰۰۶ در منطقه چابهار پرداختند. برای کشف ارتباط بین

ضروری به نظر می‌رسد. از طرفی عامل زمان در وقوع انواع بیماری‌ها نقش بسیار مهمی دارد و تغییرات عوامل محیطی و اقلیمی از زمانی به زمان دیگر بر میزان بروز بیماری‌ها و تشدید آن‌ها بسیار موثر می‌باشند. بنابراین اهداف اصلی این تحقیق، بررسی فصلی ارتباط میان متغیرهای آب و هوایی و توپوگرافی بر وقوع بیماری لپتوسپیروز و تهیه نقشه‌های فصلی آسیب‌پذیری بیماری به منظور کسب بینش عمیق به تعاملات بین شرایط آب و هوایی و محیطی با بروز بیماری لپتوسپیروز و شناسایی مناطق در معرض خطر بالای بیماری در سطح بخش‌های استان‌های شمالی ایران می‌باشد. بر این اساس این مقاله در ۶ بخش تنظیم شده است. پس از بخش مقدمه، در بخش دوم به بررسی مطالعات گذشته در زمینه بیماری‌های واگیر و مشترک بین انسان و دام، به خصوص بیماری لپتوسپیروز و استفاده از سیستم اطلاعات مکانی در مدلسازی بیماری لپتوسپیروز پرداخته شده است. در بخش سوم، منطقه مورد مطالعه و داده‌های مورد استفاده در تحقیق به همراه مفاهیم پایه و نظری روش‌های مورد استفاده در تحقیق توضیح داده شده‌اند. در بخش چهارم، نحوه مدلسازی بیماری جهت تهیه نقشه آسیب‌پذیری بیماری لپتوسپیروز و روش ارزیابی مدل بدست آمده تشریح شده‌اند. در بخش پنجم به نتایج حاصل از تحقیق به همراه بحث پرداخته شده است. در نهایت در بخش ششم، نتیجه‌گیری نهایی و پیشنهاداتی برای تحقیقات آینده ارائه شده‌اند.

۲- پیشینه تحقیق

به دلیل اهمیت موضوع بیماری در جهان از گذشته‌های دور تا به امروز پژوهش‌ها و روش‌های متعددی در زمینه بیماری‌های مختلف انجام شده‌اند. ریشه‌کنی بیماری آبله در جهان توسط گروه هندرسون از سازمان جهانی بهداشت [۱۲]، تایید نقش مخرب مصرف مواد دخانی در ایجاد بسیاری از بیماری‌ها به ویژه سرطان ریه، کشف راه‌های انتقال و روش پیشگیری از بیماری

¹ Remote Sensing

² Normalized Difference Vegetation Index

[۱۹]. در مطالعه‌ای گلشا و همکاران در سال ۲۰۰۷ میلادی اعلام نمودند که لپتوسپیروز یک بیماری شغلی محسوب می‌شود و در شالیکاران خطر ابتلا به این بیماری بیشتر است [۲۰]. مطالعات حاکی از شیوع بیماری لپتوسپیروز در مردان نسبت به زنان و نیز روستانشینان نسبت به شهرنشینان هستند و بیانگر این نکته می‌باشند که شیوع بیماری به‌طور عمده در ماه‌های گرم سال بوده و بیشتر بیماران سابقه کار در مزرعه برنج و سابقه تماس با حیوانات اهلی و آب‌های سطحی داشتند [۶، ۲۱] و همچنین رابطه سن و فعالیت در شالیزار، با ابتلا به بیماری ارتباط تنگاتنگ دارد [۲۲، ۲۳]. اما در خصوص بیماری لپتوسپیروز انسانی و استفاده از سامانه اطلاعات مکانی در تحلیل مکانی و مدلسازی بیماری، تاکنون مطالعات کمی در ایران صورت گرفته است. در مطالعه‌ای محمدی نیا و همکاران (۲۰۱۵)، به بررسی اثر پارامترهای محیطی بر بروز بیماری لپتوسپیروز با استفاده از روش‌های رگرسیون وزن‌دار جغرافیایی (GWR^1) و کمترین مربعات معمولی (OLS^2) پرداختند. نتایج به‌دست آمده از معیار ارزیابی R^2 ، به خوبی برتری روش GWR نسبت به OLS را به‌دلیل در نظر گرفتن اثر مکانی میان موارد بیماری لپتوسپیروز نشان داد [۲۴]. همچنین ایشان در مطالعه‌ای دیگر نشان دادند که بیشترین موارد وقوع بیماری در نواحی مرکزی استان گیلان به‌دلیل دارابودن زمین‌های کشاورزی در مناطق پست و با شیب کم و در ماه‌های خرداد تا شهریور اتفاق افتاده‌اند [۲۵]. تمامی این مطالعات، بر تحلیل و بررسی ارتباط میان وقوع بیماری و پارامترهای موثرش به صورت سالیانه تمرکز کرده‌اند و طبق بهترین دانش، بررسی فصلی ارتباط میان پارامترهای موثر با شیوع بیماری لپتوسپیروز و تهیه نقشه‌های فصلی آسیب‌پذیری بیماری جهت تعیین مناطق بحرانی و در

وقوع بیماری وبا و متغیرهای آب و هوایی از روش‌های سامانه اطلاعات مکانی و پایگاه داده مکانی، برای مدلسازی انتشار بیماری از روش‌های خوشه‌بندی فازی و تحلیل‌های آماری بر اساس مدل رگرسیون منطقی استفاده نمودند. نتایج این بررسی، وجود رابطه بین وقوع بیماری وبا با دما و رطوبت بالا، بارش کم و نزدیکی به مرزهای شرقی ایران بوده است [۱۶]. لاو و همکاران در سال ۲۰۱۲ میلادی، به تهیه نقشه آسیب‌پذیری بیماری لپتوسپیروز در آمریکن ساموا پرداختند. در این مطالعه از متغیرهای محیطی نظیر بارش، ارتفاع، پوشش گیاهی و ... برای مدل رگرسیون لجستیک چند متغیره استفاده نمودند. نتایج به‌دست آمده حاکی از طبقه‌بندی صحیح ۸۴ درصد از موارد گزارش شده در این مطالعه بوده است. از طرفی زندگی در ارتفاع کمتر از میانگین ارتفاع روستاها، زمین‌های گلی، مناطق کشاورزی و تراکم زیاد خوکدانی‌های منازل بیشترین تاثیر را در میزان بروز بیماری لپتوسپیروز از خود نشان دادند [۱۷]. مطالعات انجام شده در ایران در زمینه بیماری لپتوسپیروز اکثراً محدود به جنبه‌های بالینی و پزشکی بیماری بوده‌اند و بر شناخت ویژگی‌های فردی بیماران و مخازن بیماری و بیشتر در مورد نوع دامی بیماری تمرکز داشته‌اند. به‌عنوان نمونه، در مطالعه‌ای در بوشهر توسط وحدت و همکاران در سال ۲۰۰۵ میلادی، ارتباط میان بارندگی شدید و بروز لپتوسپیروز گزارش گردید. جاری شدن سیلاب پس از بارندگی برای رخداد لپتوسپیروز بسیار مناسب است، زیرا این وضعیت، مانع جذب ادرار حیوانات به خاک و یا تبخیر آن می‌شود و لپتوسپیراها به راحتی بر سطح آب‌ها روان شده، در لجن‌ها بر جای می‌مانند [۱۸]. همچنین علوی و خوشخو در سال ۲۰۱۴ میلادی، با انجام مطالعات پزشکی و تقسیم افراد به دو دسته شالیکار و غیر شالیکار دریافتند که احتمال وقوع بیماری لپتوسپیروز در شالیکاران استان خوزستان مخصوصاً جنسیت مرد و سنین بالای ۳۵ سال بسیار بیشتر از سایرین است

¹ Geographically Weighted Regression

² Ordinary Least Square

تعداد در مجموع به ترتیب فراوانی، ۷۳ درصد موارد بیماری مربوط به استان گیلان (۱۳۵۷ مورد بیماری)، ۱۸ درصد موارد بیماری مربوط به استان مازندران (۳۳۵ مورد بیماری) و ۹ درصد موارد بیماری مربوط به استان گلستان (۱۶۷ مورد بیماری) گزارش شده است. از طرفی، در سطح بخش‌های استان‌های شمالی، بیشترین موارد وقوع بیماری به ترتیب فراوانی، متعلق به رشت (۲۷/۵ درصد)، صومعه سرا (۸/۲ درصد) در استان گیلان، لاله آباد و قائمشهر (۷ درصد) در استان مازندران و گرگان (۴ درصد) در استان گرگان می‌باشد. مطابق با شکل (۱)، توزیع بیماری در نواحی مرکزی استان گیلان، مناطق شرقی استان مازندران و مناطق غربی استان گلستان بیشتر از سایر نواحی می‌باشد.

نتایج حاصل از بررسی‌های آماری مشخص نمود که روند بروز بیماری لپتوسپیروز طی ماه‌های مختلف سال تغییرات محسوسی داشته است و فراوانی به این صورت بوده است که اکثر موارد بروز بیماری از اواسط ماه اردیبهشت شروع شده و در ماه‌های خرداد تا شهریور به بالاترین میزان رسیده است و سپس کاهش یافته است. در مجموع ۶ سال، بیش‌ترین میزان فراوانی بیماری به ترتیب در فصل‌های تابستان (۶۷ درصد)، بهار (۲۳ درصد)، پاییز (۹ درصد) و زمستان (۱ درصد) بوده است. بیش‌ترین موارد وقوع بیماری در فصول بهار و تابستان می‌تواند احتمالاً مربوط به شروع فعالیت شالیکاران و کشاورزان، افزایش رطوبت، دما و بارندگی شدید باشد (شکل (۲)).

معرض خطر بالای بیماری لپتوسپیروز در هیچ کدام یک از مطالعات پیشین، صورت نگرفته است. همچنین در هیچ کدام از پژوهش‌های انجام شده، از روش یادگیری ماشینی ماشین بردار پشتیبان (SVM^۱) با قدرت طبقه‌بندی داده‌ها در کلاس‌های مختلف، به‌عنوان روشی برای تهیه نقشه آسیب‌پذیری بیماری‌ها استفاده نشده است.

۳- مواد و روش‌ها

در این قسمت از پژوهش علاوه بر اشاره به منطقه مورد مطالعه و داده‌های جمع‌آوری شده جهت تهیه نقشه آسیب‌پذیری بیماری لپتوسپیروز، مفاهیم پایه و نظری روش‌های مورد استفاده در تحقیق شامل تحلیل همبستگی پیرسون^۲ و ماشین بردار پشتیبان تشریح شده‌اند.

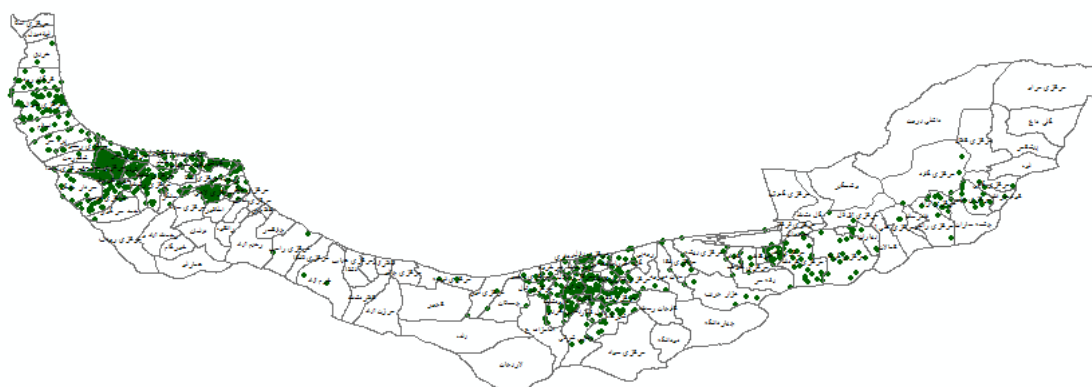
۳-۱- منطقه مورد مطالعه

استان‌های حاشیه دریای خزر شامل استان‌های گیلان، مازندران و گلستان، با جمعیتی معادل ۷,۳۳۱,۸۳۱ نفر طبق سرشماری سال ۱۳۹۰ به عنوان مناطق مورد مطالعه در این تحقیق انتخاب شده‌اند. این استانها مساحتی حدود ۵۸.۲۵۰ کیلومتر مربع واقع در شمال ایران را می‌پوشانند و از نظر تقسیمات سیاسی، در مجموع متشکل از ۱۱۹ بخش^۳ می‌باشند. با توجه به موقعیت جغرافیایی و شرایط آب و هوایی استان‌های مورد مطالعه و نتایج حاصل از ایستگاه‌های هواشناسی، این استان‌ها دارای آب و هوای معتدل و مرطوب و از جمله استان‌های بارش خیز می‌باشند. در این تحقیق در فاصله زمانی فروردین ماه سال ۱۳۸۸ تا اسفند ماه سال ۱۳۹۳ (۷۲ ماه) در مجموع ۱۸۵۹ مورد بیماری لپتوسپیروز در سطح بخش‌های استان‌های گیلان، مازندران و گلستان ثبت و ضبط شده است. از این

^۱ Support Vector Machine (SVM)

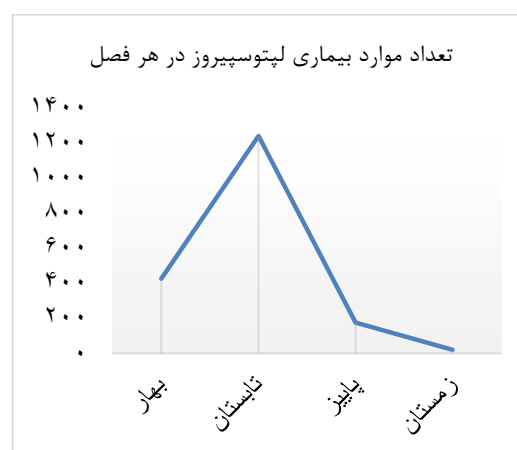
^۲ Pearson's correlation coefficient

^۳ District



شکل ۱: پراکندگی موارد وقوع بیماری لپتوسپیروز در بخش‌های استان‌های شمالی ایران در سال‌های ۱۳۸۸-۱۳۹۳

لپتوسپیروز مربوط به جنسیت مرد بوده، اما این بیماری برای هر دو جنسیت زن و مرد تهدید کننده است. از طرفی، بروز این بیماری در تمامی گروه سنی مشاهده شده است که از میان آن‌ها بیشترین بروز بیماری مربوط به جوانان و میانسالان بوده است. علت چنین نتایجی ممکن است به دلیل پوشش کمتر و تعاملات اجتماعی و محیطی بیشتر مردان نسبت به زنان و بیانگر خطر مواجهه در مشاغل باشد که مردان بیشتر در آن حضور دارند، به همین دلیل بیماری بین مردان جوان و میانسال که سنین فعال کاری می‌باشند، از پسران و سالخوردگان شیوع بیشتری دارد (جدول (۱)).



شکل ۲: فصول وقوع بیماری لپتوسپیروز (۱۳۸۸-۱۳۹۳)

همچنین بررسی‌های آماری مشخص نمود که اگرچه در تمام بخش‌های استان‌های شمالی ایران در هر ۶ سال مورد مطالعه، بیش‌ترین فراوانی موارد بروز بیماری

جدول ۱: خصوصیات کلی بیماران مبتلا به لپتوسپیروز در استان‌های شمالی ایران در سال‌های ۱۳۸۸ تا ۱۳۹۳

سال	ویژگی‌های جمعیت شناختی												
	تعداد موارد وقوع بیماری در هر فصل				گروه سنی							جنسیت	
	بهار	تابستان	پاییز	زمستان	۰-۴	۵-۱۰	۱۱-۱۴	۱۵-۲۴	۲۵-۳۴	۳۵-۵۵	۵۶+	مرد	زن
۱۳۸۸	۴۹	۲۱۳	۳۳	۰	۱	۲	۱	۳۳	۵۳	۱۷۸	۲۷	۲۱۳	۸۲
۱۳۸۹	۱۱۰	۴۲۰	۳۰	۳	۰	۰	۱	۵۶	۷۸	۳۵۹	۶۹	۴۱۵	۱۴۸
۱۳۹۰	۴۲	۱۶۰	۲۴	۵	۰	۰	۰	۱۳	۴۵	۱۳۷	۳۶	۱۵۷	۷۴
۱۳۹۱	۸۰	۱۲۱	۲۹	۳	۰	۰	۲	۱۹	۴۱	۱۳۷	۳۴	۱۸۴	۴۹
۱۳۹۲	۶۰	۱۶۳	۳۴	۶	۰	۰	۳	۲۲	۴۴	۱۷۰	۲۴	۱۷۴	۸۹
۱۳۹۳	۸۵	۱۶۱	۲۵	۳	۰	۰	۱	۱۶	۴۶	۱۷۴	۳۷	۱۸۳	۹۱
مجموع	۴۲۶	۱۲۳۸	۱۷۵	۲۰	۱	۲	۸	۱۵۹	۳۰۷	۱۱۵۵	۲۲۷	۱۳۲۶	۵۳۳

۳-۲- جمع آوری و آماده سازی داده ها

در این پژوهش اطلاعات مربوط به میزان شیوع بیماری لپتوسپیروز طی سال های ۱۳۸۸-۱۳۹۳ از مرکز مدیریت بیماری های واگیر وزارت بهداشت به تفکیک سن، جنس، تاریخ بروز بیماری و آدرس محل سکونت بیمار به صورت ماهانه جمع آوری شدند. همچنین داده های جمعیتی طی این سال ها از مرکز آمار ایران به دست آمده اند. سپس با استفاده از داده های جمعیتی و وقوع بیماری، نرخ شیوع استاندارد^۱ بیماری لپتوسپیروز طبق رابطه (۱)، محاسبه گردید.

$$\text{نرخ شیوع} = \frac{\text{تعداد افراد مبتلا در یک جمعیت در یک زمان مشخص}}{\text{تعداد کل جمعیت در همان زمان}} * 100000$$

همچنین به دلیل تاثیر گذاری بالایی که عوامل محیطی بر شیوع بیماری ها دارند، لذا متغیرهای پوشش گیاهی، توپوگرافی (شامل ارتفاع و شیب)، و آب و هوایی (از قبیل میانگین دما، بارش، میانگین رطوبت و تعداد روزهای یخبندان) که بیشترین استفاده را در مطالعات محیطی دارند [۱۴، ۱۷، ۲۶، ۲۷]، استفاده شده اند. به منظور بررسی اثر پوشش گیاهی بر روی شیوع بیماری لپتوسپیروز از شاخص پوشش گیاهی نرمال شده استخراج شده از تصاویر ماهواره ای مادیس (MODIS) استفاده شده است. این تصاویر با قدرت تفکیک مکانی ۲۵۰ متر در بازه های زمانی ۱۶ روزه دریافت شده اند. همچنین مدل رقومی ارتفاعی استخراج شده از تصاویر ماهواره ای ماموریت مکان نگاری شاتل رادار (SRTM^۲) با قدرت تفکیک ۳۰ متر به منظور تولید نقشه های ارتفاعی و شیب مورد استفاده قرار گرفت. داده های هواشناسی شامل میانگین دمای هوا، مجموع بارش فصلی، میانگین رطوبت هوا و تعداد روزهای یخبندان به صورت ماهانه

از سازمان هواشناسی ایران در طی بازه زمانی مورد نظر جمع آوری شده اند. با توجه به اینکه داده های هواشناسی به صورت نقطه ای وجود دارند. لذا از روش درونیایی وزندهی معکوس فاصله (IDW^۳) برای داشتن سطحی پیوسته استفاده شده است. با استفاده از روش IDW هر لایه نقطه ای به یک لایه رستری تبدیل شده و لایه رستری مربوطه براساس محدوده منطقه مورد مطالعه برش داده شده است. سپس برای هر پلی گون متوسط پیکسل های داخل پلی گون، به عنوان پارامتر آن پلی گون در نظر گرفته شده است. در نهایت، بعد از جمع آوری داده های مورد نیاز، به منظور یکسان نمودن دامنه مقادیر داده ها و به حداقل رساندن اثر مقیاس، مقادیر داده ها در بازه [۰-۱] نرمال سازی شده اند. سپس به منظور به کارگیری هر کدام از داده ها در مدل سازی، لایه های اطلاعاتی مورد نیاز اعم از نقشه های ارتفاعی، شیب، پوشش گیاهی، میانگین دمای هوا، میانگین رطوبت هوا، مجموع بارش سالیانه، تعداد روزهای یخبندان و نرخ شیوع بیماری لپتوسپیروز در فصول گذشته به فرمت شیپ فایل در نرم افزار ArcGIS 10.2.2 ایجاد شدند.

۳-۳- تحلیل همبستگی پیرسون

رابطه همبستگی پیرسون به بررسی ارتباط میان دو یا چند متغیر پرداخته و ضریب آن را محاسبه می نماید. همبستگی بین متغیرها ممکن است مثبت و یا منفی باشد [۲۸]. عموماً متداول ترین روش برای بیان همبستگی میان متغیرهای دارای اندازه های مقیاس فاصله ای و نسبی، استفاده از ضریب همبستگی پیرسون می باشد که در این تحقیق از آن استفاده شده است و از رابطه (۲) محاسبه می شود [۲۹]:

$$r_{x,y} = \frac{\sum x*y}{N*S_x*S_y} \quad \text{رابطه (۲)}$$

در رابطه فوق، $r_{x,y}$ مقدار همبستگی پیرسون بین

^۱ Standardized Incidence Rate (SIR)

^۲ Shuttle Radar Topography Mission (SRTM)

^۳ Inverse distance weighting

با استفاده از نتایج تابع PCA، فرض می‌شود که متغیرهای انتخابی مستقل از یکدیگر می‌باشند و رابطه هم‌خطی میان آنها برقرار نیست. با توجه به اینکه در این تحقیق، همبستگی بین متغیرهای توپوگرافی و محیطی کمتر از ۰/۸ بوده است، لذا هیچ کدام از متغیرها از فرآیند مدلسازی بیماری حذف نشدند و از تمامی متغیرها در مدلسازی بیماری لپتوسپیروز به‌عنوان متغیرهای مستقل استفاده شده است. تابع PCA، از روش‌های آماری چند متغیره است که می‌توان از آن برای کاهش تعداد متغیرها و تفسیر بهتر اطلاعات استفاده نمود. هدف از تحلیل مولفه‌های اصلی آن است که واریانس موجود در داده‌های چند متغیره را به مولفه‌هایی تجزیه کند که اولین مولفه تا آنجا که ممکن است علت بیشترین واریانس موجود در داده‌ها باشد. دومین مولفه علت بیشترین واریانس ممکن بعد از مولفه اول باشد و همچنین برای بقیه مولفه‌ها، تا جایکه هیچگونه همبستگی بین مولفه‌ها موجود نباشد. با اعمال این روش، متغیرهای ورودی اولیه به مولفه‌های جدید و بدون همبستگی تبدیل می‌شوند؛ به طوری که مولفه‌های ایجاد شده ترکیبی خطی از متغیرهای ورودی اند [۳۲].

۳-۴- ماشین بردار پشتیبان

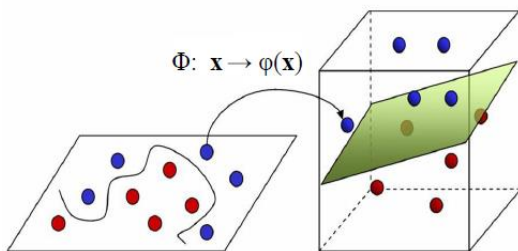
ماشین بردار پشتیبان به‌عنوان یک روش موفق طبقه‌بندی در میان روش‌های یادگیری ماشینی معرفی شده است و از جمله روش‌های نسبتاً جدیدی است که در سال‌های اخیر کارایی خوبی نسبت به روش‌های قدیمی‌تر برای طبقه‌بندی از جمله شبکه‌های عصبی از خود نشان داده است [۳۳]. این روش یک طبقه‌بندی کننده نظارت‌شده و دودویی^۲ است که علاوه بر کلاسه‌بندی و تشخیص الگوها، برای پیش‌بینی و حل مسائل پیچیده نیز مورد استفاده قرار می‌گیرد [۳۴]. ایده اصلی در SVM این است که با فرض

متغیرهای x و y ، N : تعداد آزمودنی‌ها، $\sum x * y$: مجموع حاصل ضرب تفاضل مقادیر از میانگین، Sx : انحراف استاندارد مقادیر x ، Sy : انحراف استاندارد مقادیر y می‌باشد. مقدار ضریب همبستگی پیرسون بین -1 تا $+1$ تغییر می‌کند. اگر مقدار ضریب برابر با $+1$ باشد، همبستگی مثبت کامل را نشان می‌دهد و اگر برابر با -1 باشد، نشان‌دهنده همبستگی کامل و منفی است. در حالت کلی، مقدار ضریب همبستگی در بازه $(-1, +1)$ بیانگر برقراری همبستگی ناقص میان دو متغیر می‌باشد. زمانیکه مقدار ضریب برابر صفر باشد، نشان می‌دهد که هیچ رابطه‌ای بین دو متغیر وجود ندارد [۳۰]. در این تحقیق از نرم‌افزار SPSS جهت انجام تحلیل همبستگی پیرسون بهره گرفته شده است. در این راستا، از متغیرهای توپوگرافی و محیطی مانند ارتفاع، شیب، پوشش گیاهی، متوسط رطوبت هوا، متوسط دمای هوا، مجموع بارش فصلی، تعداد روزهای یخبندان و نرخ شیوع بیماری لپتوسپیروز در فصول گذشته، به عنوان پارامترهای مستقل استفاده نموده و به بررسی همبستگی بین این پارامترها و شیوع بیماری لپتوسپیروز در فصل مورد نظر به عنوان پارامتر وابسته، تعیین اهمیت و نوع تأثیر هر یک از آن‌ها و مشخص نمودن متغیرهایی که بیشترین تأثیر را بر وقوع بیماری دارند با استفاده از تحلیل همبستگی پیرسون پرداخته شده است (لازم به ذکر است که به دلیل تعداد خیلی کم موارد بیماری در فصل زمستان، همبستگی بین متغیرها با بروز بیماری لپتوسپیروز در فصول بهار، تابستان و پاییز بررسی شده است). در ضمن قبل از انجام تجزیه و تحلیل، میزان همبستگی بین متغیرها با استفاده از تابع تحلیل مولفه‌های اصلی (PCA^۱) بررسی شده است تا متغیرهایی که کمترین همبستگی (همبستگی کمتر از ۰/۸) را با یکدیگر دارند انتخاب شوند [۳۱]. در نتیجه

^۲ Binary

^۱ Principal Component Analysis

یکی از مزایای ماشین بردار پشتیبان محدود نبودن آن به فضای تصمیم‌گیر خطی است. در حالی که بسیاری از پدیده‌های طبیعی قابل مدل‌سازی در فضاهای خطی نمی‌باشند. از این‌رو، ماشین بردار پشتیبان در مسائلی که داده‌ها به صورت خطی جداپذیر نیستند، با استفاده از هسته‌های غیرخطی که اصطلاحاً توابع کرنل^۴ نامیده می‌شوند، داده‌ها را به فضایی با ابعاد بیشتر نگاشت می‌دهد تا بتوان آن‌ها را در این فضای جدید به صورت خطی جدا نمود [۳۶] (شکل ۵).

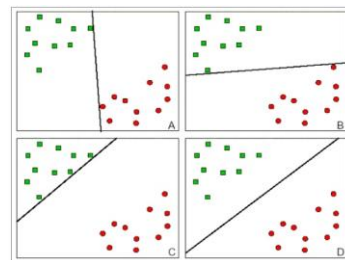


شکل ۵: استفاده از تابع کرنل برای نگاشت به فضایی با

ابعاد بالاتر با قابلیت جداسازی خطی داده‌ها [۳۶]

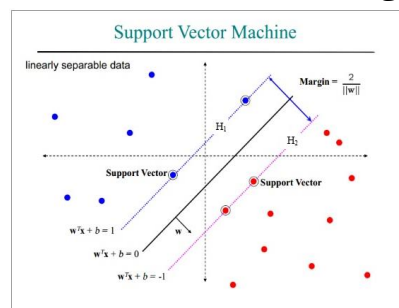
برای تعمیم SVM به حالت چندکلاسه (طبقه‌بندی چندگانه)، روش‌های مختلفی از جمله روش یک در برابر همه^۵ و روش یک در برابر یک^۶ معرفی شده‌اند [۳۷]. در روش یک در برابر همه برای طبقه‌بندی i امین کلاس، فرض می‌شود که نقاط موجود در کلاس i به‌عنوان نمونه‌های مثبت و تمامی نقاط موجود در کلاس‌های دیگر به‌عنوان نمونه‌های منفی در نظر گرفته شوند. زمانی که المان‌های این کلاس مشخص شدند. کلاس بعدی را در نظر گرفته و مجدداً مراحل تکرار می‌شود. این کار تا زمانی که تمام کلاس‌ها طبقه‌بندی شوند، ادامه پیدا می‌کند. بنابراین در این روش، برای N کلاس موجود، N طبقه‌بندی کننده ایجاد می‌شود که هر یک از آنها کلاس مربوط به خود را از سایر کلاس‌ها جدا می‌کنند [۳۸]. اما در روش یک در برابر یک، الگوریتم به ازای هر جفت کلاس موجود

جداپذیری کلاس‌ها از هم، ابرصفحاتی^۱ که قادر به جدانمودن کلاس‌ها از هم باشند را به دست می‌آورد و دو کلاس را با استفاده از یک مرز خطی از هم جدا می‌کند [۳۵]. مطابق با شکل (۳)، بی‌نهایت خط جداکننده می‌توان ترسیم نمود که نمونه‌های دو کلاس را از یکدیگر جدا کند. بنابراین در تئوری SVM، هدف تعیین خط جداکننده بهینه از میان بی‌نهایت خطوط جداکننده می‌باشد.



شکل ۳: مرزهای مختلف در جداسازی دو کلاس [۳۵]

در تقسیم خطی، سعی بر این است خطی انتخاب شود که بیشترین فاصله را از نمونه‌های دو کلاس و در نتیجه خطای تعمیم کمتری داشته باشد. این فاصله در تئوری SVM، حاشیه^۲ نامیده می‌شود. در واقع ماشین بردار پشتیبان ابر صفحه‌ای ایجاد می‌کند که کلاس‌ها را به گونه‌ای از هم جدا کند تا فاصله میان نزدیکترین نمونه‌های دو کلاس در راستای عمود بر مرز تصمیم‌گیری بیشینه شود. نزدیکترین نمونه‌ها به ابرصفحه جداکننده بهینه، بردار پشتیبان^۳ نامیده می‌شود [۳۳] (شکل ۴).



شکل ۴: ابرصفحه جداکننده خطی در فضای خطی [۳۵]

^۴ Kernel Function

^۵ One against all

^۶ One against one

^۱ Hyperplane

^۲ Margin

^۳ Support Vector

بهینه مدل، از روش اعتبار سنجی متقابل^۱ با به کارگیری ۳۰ درصد از داده‌های آموزشی استفاده شده است (شکل (۶) الف). سرانجام مدل آموزش دیده، با استفاده از ۲۰ درصد باقی مانده از کل داده‌ها، به منظور ارزیابی دقت مدل و تهیه نقشه آسیب پذیری بیماری لپتوسپیروز برای فصول سال ۹۳، مورد استفاده قرار گرفته شده است (شکل (۶) ب). همچنین در این تحقیق، با توجه به تعداد کم کلاس‌ها (۵ کلاس)، از روش طبقه بندی چندگانه یک در برابر همه استفاده شده است. بنابراین برای مسئله بیماری با ۵ کلاس نرخ شیوع بیماری، ۵ SVM دودویی ایجاد شده است.

در این پژوهش به منظور ارزیابی مدل SVM از منحنی مشخصه عملکرد سیستم^۲ (ROC) استفاده شده است. منحنی ROC ابزاری است که عمدتاً برای تعیین میزان کارایی روش‌های تشخیصی و الگوریتم‌های دسته بندی به کار می‌رود [۴۱]. در واقع، یک نمودار ROC، نمایش گرافیکی سبک و سنگین کردن بین منفی و مثبت کاذب است که برای هر پیکسل در منطقه مورد مطالعه یک منحنی به طور تصادفی مشخص می‌شود [۴۲]. سطح زیر این منحنی (AUC^3) برابر با احتمال قدرت تشخیص میان نقاط وجود و عدم وجود توسط یک مدل است. مقادیر مختلف سطح زیر منحنی بین ۰٫۵ تا ۱ است. چنانچه سطح زیر منحنی برابر با ۰٫۵ باشد بیان کننده تصادفی بودن مدل بوده و اگر مقدار برابر با ۱ باشد، مدل به بهترین نحو می‌تواند نقاط وجود و عدم وجود را از یکدیگر تفکیک نماید. سطح زیر منحنی بین ۰٫۷ تا ۰٫۸ بیانگر یک مدل خوب، مقادیر بین ۰٫۸ تا ۰٫۹ مدل عالی و سطح زیر منحنی بیش از ۰٫۹ نشان دهنده قدرت تشخیص بسیار عالی مدل می‌باشد [۴۳].

آموزش داده می‌شود. بنابراین برای N کلاس موجود، به تعداد $N(N-1)$ طبقه بندی کننده ایجاد می‌شود که هر طبقه بندی کننده وظیفه طبقه بندی یک جفت کلاس (i,j) را بر عهده دارد. به طوریکه i نمونه‌های مثبت و j نمونه‌های منفی در نظر گرفته می‌شوند [۳۹]. اینکه کدامیک از این دو روش برای کاربردی خاص مناسب تر است، بسته به تعداد کلاس‌ها و همچنین تعداد نمونه‌های آموزشی دارد و برای کاربردهای مختلف می‌تواند متفاوت باشد. اما به طور کلی، زمانی که تعداد کلاس‌ها کم باشد روش یک در برابر همه روش مناسب تری است. اما زمانی که تعداد کلاس‌ها بسیار زیاد است، این روش دقت مناسبی را حاصل نمی‌کند. در این شرایط روش یک در برابر یک می‌تواند نتایج دقیق تری را حاصل نماید. همچنین روش یک در برابر یک، سرعت بالاتری نسبت به روش یک در برابر همه دارد. بنابراین در مسائلی با داده‌های زیاد که زمان محاسبات از مشکلات محدود کننده می‌باشد، روش یک در برابر یک به عنوان روش مناسب تر پیشنهاد می‌شود [۴۰].

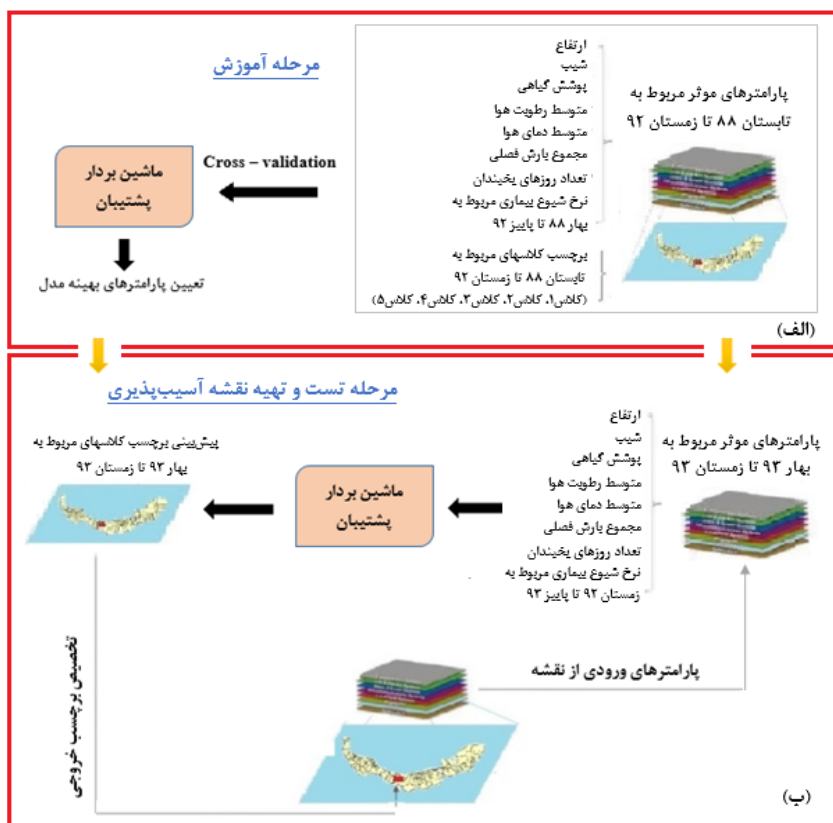
۳-۵- مدل سازی و ارزیابی

در این تحقیق، به منظور آموزش مدل SVM، ابتدا داده‌های مربوط به وقوع بیماری لپتوسپیروز در فصول هر ۶ سال مورد بررسی، به ۵ کلاس مختلف با برجسب‌های ۱ تا ۵ تقسیم شدند (نحوه تعریف ۵ کلاس بدین صورت بوده است که داده‌های در محدوده ۰٫۲-۰٫۴ در کلاس ۱، داده‌های در محدوده ۰٫۴-۰٫۶ در کلاس ۲، داده‌های در محدوده ۰٫۶-۰٫۸ در کلاس ۳ و در نهایت داده‌های در محدوده ۰٫۸-۱ در کلاس ۵ طبقه بندی شدند). سپس پارامترهای موثر در وقوع بیماری به همراه برجسب‌های کلاس وقوع بیماری مربوط به فصول سال‌های ۹۲-۸۸ که ۸۰ درصد از کل داده‌ها را شامل می‌شوند، به مدل SVM تغذیه شدند. در مرحله آموزش مدل، به منظور تعیین پارامترهای

¹ Cross- validation

² Receiver Operating Characteristic (ROC)

³ Area Under the Curve (AUC)



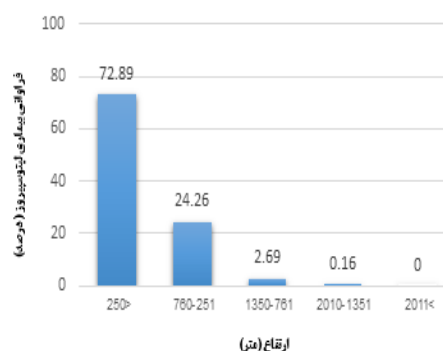
شکل ۶: فرآیند آموزش مدل SVM

۴- نتایج و بحث

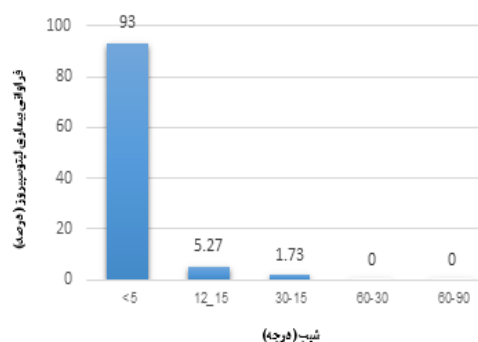
مقایسه محل وقوع موارد بیماری لپتوسپیروز در ۶ سال مورد مطالعه با مدل رقومی زمین نشان می‌دهد که بیش از ۹۵ درصد از موارد بروز بیماری لپتوسپیروز در ارتفاعات کمتر از ۷۶۰ متر از سطح دریا قرار گرفته است و کمتر از ۵ درصد از موارد گزارش شده در ارتفاعات بالاتر از ۷۶۰ متر از سطح دریا واقع شده‌اند (شکل ۷). همچنین بررسی‌های انجام شده در رابطه با تاثیر شیب زمین در بروز بیماری لپتوسپیروز، حاکی از آن است که بیش از ۹۷ درصد از موارد گزارش شده بیماری لپتوسپیروز در شیب‌های کمتر از ۱۵ درجه اتفاق افتاده‌اند. علاوه بر این، بر اساس نتایج آنالیز همبستگی پیرسون (جدول ۲)، همبستگی منفی بین ارتفاع و شیب با بروز بیماری لپتوسپیروز وجود دارد و تراکم بیماری لپتوسپیروز با طبقه‌های ارتفاع و شیب در هر فصل به‌طور جداگانه ارتباط معنی داری را

نشان می‌دهد و نشان دهنده این مطلب است که بیشترین موارد بیماری در مناطق با ارتفاع کم و شیب پائین یافت شده‌اند، به طوری که با افزایش ارتفاع و شیب زمین احتمال وقوع بیماری لپتوسپیروز رو به کاهش بوده است. این امر با این واقعیت که اغلب زمین‌های کشاورزی و شالیکاری در ارتفاعات قرار نگرفته‌اند و بیشتر در مناطق پست با شیب کم حضور دارند، منطبق است. همچنین در این مناطق، به دلیل ارتفاع و شیب پائین، تجمع آب‌های سطحی و راکد افزایش یافته و مانع از جذب عفونت بیماری در خاک و یا تبخیر آن می‌شود و در نتیجه باعث می‌شود که احتمال وقوع بیماری در انسان‌ها به دلیل افزایش سطح تماس با عفونت بیماری، بیشتر شود (شکل ۸).

همبستگی ضعیف بین این متغیرها و بروز بیماری لپتوسپیروز وجود دارد. همبستگی منفی میان وقوع بیماری و تعداد روزهای یخبندان می‌تواند به این علت باشد که در آب و هوای سرد و یخبندان، احتمال ادامه حیات، تکثیر و انتقال باکتری و عفونت بیماری کاهش می‌یابد، بنابراین از تعداد موارد وقوع بیماری تحت این شرایط آب و هوایی کاسته می‌شود. همچنین همبستگی منفی میان پوشش گیاهی و وقوع بیماری را می‌توان این‌طور تفسیر نمود که محل سکونت انسان‌ها معمولاً در مناطق با پوشش گیاهی متوسط و کم و حتی بدون پوشش گیاهی (مناطق شهری) واقع می‌باشد و همچنین فعالیت‌های شغلی از جمله کشاورزی و شالیکاری غالباً واقع در مناطق با پوشش گیاهی متراکم نمی‌باشند، بنابراین سطح تماس افراد با مخازن و عفونت بیماری کاهش یافته و در نتیجه از تعداد موارد بیماری در مناطق با پوشش گیاهی متراکم کاسته می‌شود. علاوه بر این، نرخ شیوع بیماری در فصل گذشته ضعیف‌ترین همبستگی مثبت را نشان می‌دهد که با نتایج حاصل از مطالعه انجام شده توسط بابامحمودی و همکاران (۲۰۰۶) مبنی بر اینکه انتقال بیماری لپتوسپیروز از انسان به انسان به ندرت اتفاق می‌افتد [۳]، مطابقت دارد. همچنین مقایسه محل وقوع موارد بیماری لپتوسپیروز با شرایط آب و هوایی و پوشش گیاهی نشان می‌دهد که هیچ خوشه‌ای در مناطق آب و هوای سرد و کوهستانی و مقادیر NDVI بالاتر موجود نبوده است. در حالی که مناطق با موارد وقوع با ریسک بالاتر در مناطقی با آب و هوای گرم و مرطوب به همراه بارش شدید باران و مقادیر پائین NDVI واقع شده‌اند. این شرایط، محیط مساعدی را برای تکثیر و تولید مثل مخازن و ناقلین بیماری لپتوسپیروز فراهم می‌کند. نتایج بررسی ارتباط میان متغیرهای آب و هوایی و پوشش گیاهی با بروز بیماری لپتوسپیروز در شکل (۹) و جدول (۲) ارائه شده است.

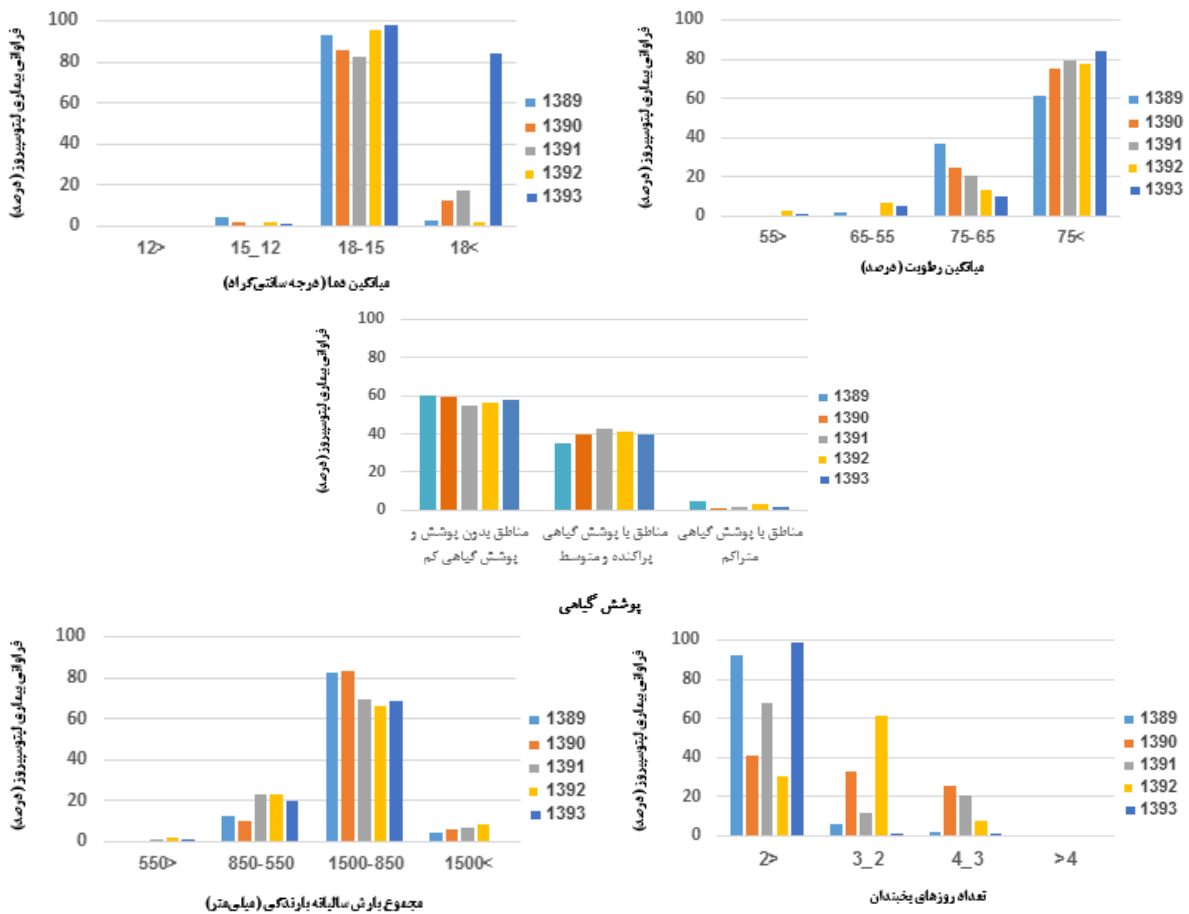


شکل ۷: فراوانی نسبی (درصد) بیماری لپتوسپیروز در خوشه‌های مکانی شناسایی شده در بازه‌های ارتفاعی مختلف در استان‌های شمالی ایران (۱۳۸۸-۱۳۹۳)



شکل ۸: فراوانی نسبی (درصد) بیماری لپتوسپیروز در خوشه‌های مکانی شناسایی شده در بازه‌های مختلف شیب زمین در استان‌های شمالی ایران (۱۳۸۸-۱۳۹۳)

همچنین نتایج آنالیز همبستگی پیرسون، حاکی از آن است که بین عناصر متوسط دمای هوا، متوسط رطوبت هوا و میزان بارش با بروز بیماری لپتوسپیروز در بخش‌های استان‌های شمالی، ارتباط معنی‌داری وجود داشته و یک همبستگی مثبت قوی بین این متغیرها و بروز بیماری لپتوسپیروز وجود دارد. این بدان معنی است که با افزایش رطوبت، دما و بارش، میزان رخداد بیماری نیز افزایش می‌یابد. بنابراین در ماه‌های گرم و مرطوب سال که همراه با بارش‌های شدید ناشی از بخار آب‌های سطحی می‌باشند، موارد وقوع بیماری بیش‌تر مشاهده می‌شود. در حالی که همبستگی بین متغیرهای پوشش گیاهی و تعداد روزهای یخبندان معنی‌دار اما منفی می‌باشد و یک



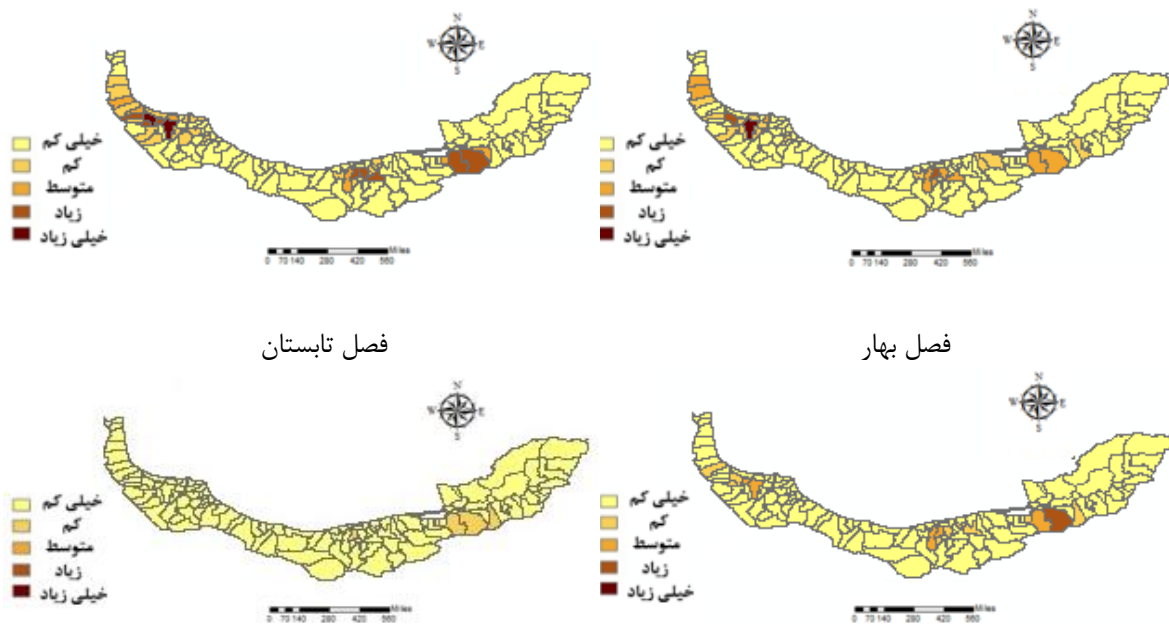
شکل ۹: فراوانی نسبی (درصد) بیماری لپتوسپیروز در طبقات مختلف پرامترهای محیطی در استان‌های شمالی ایران طی سال‌های ۱۳۸۹-۱۳۹۳

جدول ۲: نتایج آنالیز همبستگی پیرسون بین متغیرهای محیطی و توپوگرافی با بروز بیماری لپتوسپیروز در سطح بخش‌های مختلف استان‌های شمالی ایران برای فصول مختلف.

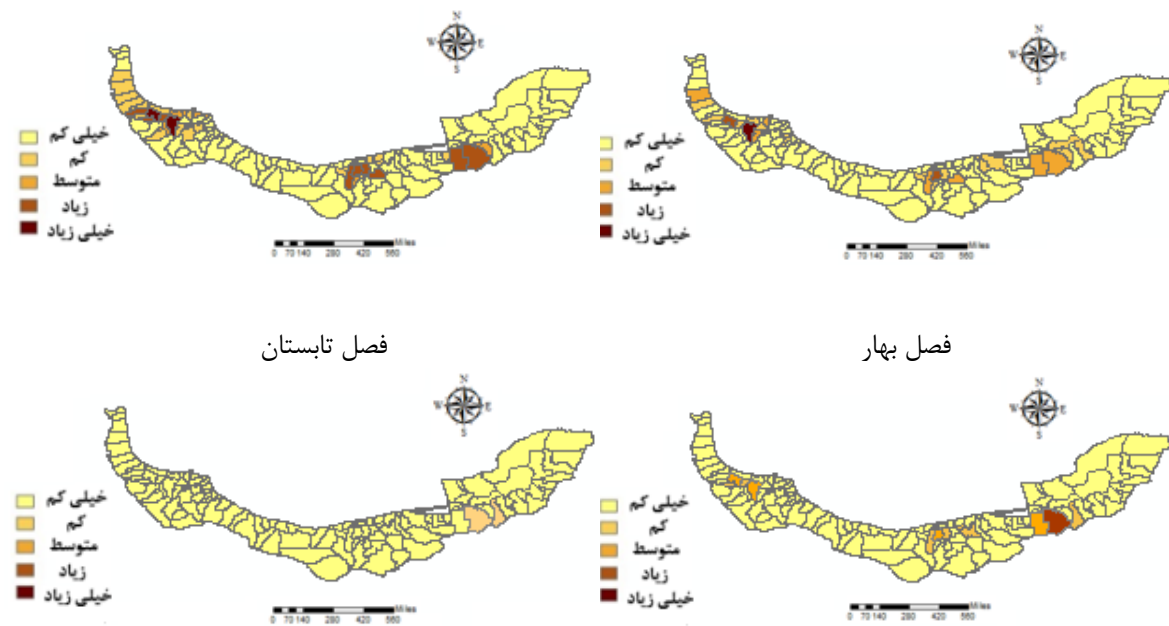
فصل / متغیر	ارتفاع	شیب	پوشش گیاهی	متوسط دما	متوسط رطوبت	مجموع بارش فصلی	تعداد روزهای یخبندان	نرخ شیوع فصل گذشته
بهار	۰٫۳۹۷	۰٫۲۶۴	۰٫۲۲۱	۰٫۶۴۸	۰٫۷۵۵	۰٫۶۱۲	۰٫۱۰۳	۰٫۰۷۹
تابستان	۰٫۴۲۹	۰٫۳۰۲	۰٫۱۸۲	۰٫۷۰۴	۰٫۷۹۱	۰٫۶۸۸	۰٫۰۹۴	۰٫۰۸۸
پاییز	۰٫۳۷۴	۰٫۲۷۶	۰٫۲۵۳	۰٫۶۰۱	۰٫۶۵۱	۰٫۵۹۳	۰٫۱۱۸	۰٫۰۶۳

نقشه‌های واقعی و آسیب‌پذیری بیماری لپتوسپیروز حاصل از مدل SVM برای هر ۴ فصل مربوط به سال ۱۳۹۳ در شکل (۱۰ و ۱۱) و نتایج ارزیابی نقشه‌های

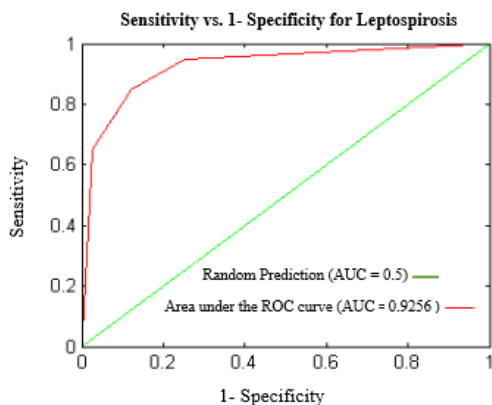
تولید شده در شکل (۱۲) و جدول (۳) نشان داده شده‌اند.



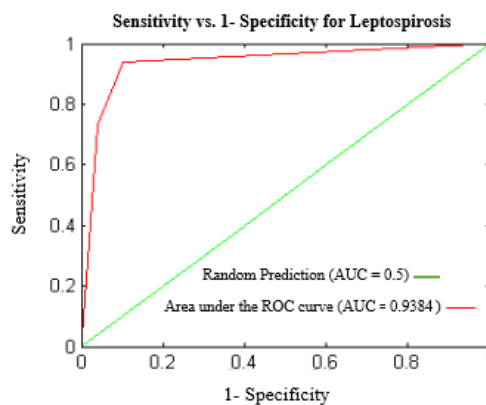
شکل ۱۰: نقشه‌های واقعی برای هر ۴ فصل مربوط به سال ۱۳۹۳.



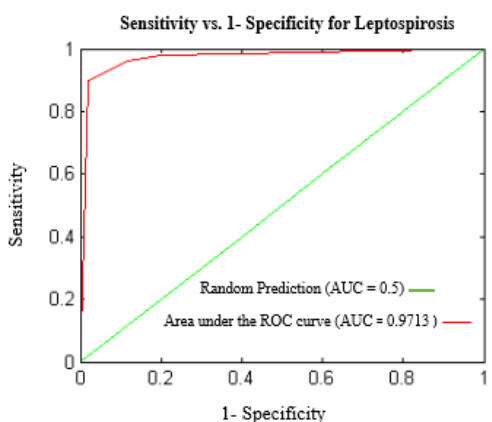
شکل ۱۱: نقشه‌های آسیب‌پذیری تهیه شده توسط مدل SVM برای هر ۴ فصل مربوط به سال ۱۳۹۳



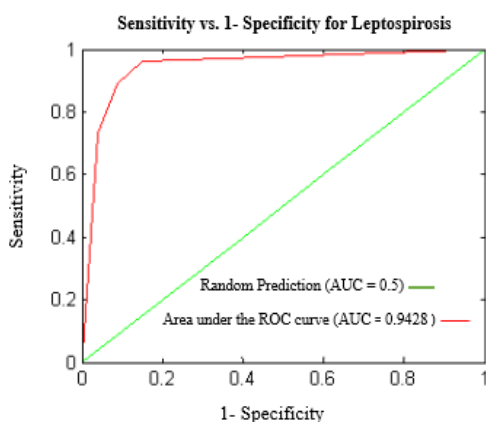
فصل تابستان



فصل بهار



فصل زمستان



فصل پاییز

شکل ۱۲: منحنی‌های ROC مربوط به نتایج مدل SVM برای هر فصل سال ۱۳۹۳.

جدول ۳: نتایج ارزیابی حاصل از مدل SVM

فصل	بهار	تابستان	پاییز	زمستان
مساحت زیر منحنی ROC حاصل از مدل SVM	۰٫۹۳۸۴	۰٫۹۲۵۶	۰٫۹۴۲۸	۰٫۹۷۱۳

آسیب‌پذیری تهیه شده در هر فصل به همراه درصد صحیح بیماری لپتوسپیروز در هر کدام از این کلاس‌ها محاسبه شدند (جدول (۴، ۵، ۶، ۷ و ۸)).

به‌منظور ارزیابی مکانی دقت مدل SVM، بدین صورت عمل شده است که تعداد موارد درست و غلط پیش بینی شده در کلاس‌های مختلف نقشه‌های

جدول ۴: دقت مکانی مدل SVM در کلاس خیلی کم

فصل	تعداد موارد در نقشه واقعی	تعداد موارد در نقشه آسیب‌پذیری تهیه شده	تعداد موارد غلط	درصد صحیح
بهار	۹۳	۹۷	۴	۹۶
تابستان	۸۴	۸۸	۴	۹۵
پاییز	۱۰۴	۱۰۸	۴	۹۶
زمستان	۱۱۵	۱۱۷	۲	۹۸

جدول ۵: دقت مکانی مدل SVM در کلاس کم

فصل	تعداد موارد در نقشه واقعی	تعداد موارد در نقشه آسیب پذیری تهیه شده	تعداد موارد غلط	درصد صحیح
بهار	۱۴	۱۲	۲	۸۶
تابستان	۱۹	۱۶	۳	۸۴
پاییز	۹	۶	۳	۶۷
زمستان	۴	۲	۲	۵۰

جدول ۶: دقت مکانی مدل SVM در کلاس متوسط

فصل	تعداد موارد در نقشه واقعی	تعداد موارد در نقشه آسیب پذیری تهیه شده	تعداد موارد غلط	درصد صحیح
بهار	۹	۷	۲	۷۸
تابستان	۶	۴	۲	۶۷
پاییز	۵	۴	۱	۸۰
زمستان	۰	۰	۰	۱۰۰

جدول ۷: دقت مکانی مدل SVM در کلاس زیاد

فصل	تعداد موارد در نقشه واقعی	تعداد موارد در نقشه آسیب پذیری تهیه شده	تعداد موارد غلط	درصد صحیح
بهار	۲	۲	۰	۱۰۰
تابستان	۸	۹	۱	۸۹
پاییز	۱	۱	۰	۱۰۰
زمستان	۰	۰	۰	۱۰۰

جدول ۸: دقت مکانی مدل SVM در کلاس خیلی زیاد

فصل	تعداد موارد در نقشه واقعی	تعداد موارد در نقشه آسیب پذیری تهیه شده	تعداد موارد غلط	درصد صحیح
بهار	۱	۱	۰	۱۰۰
تابستان	۲	۲	۰	۱۰۰
پاییز	۰	۰	۰	۱۰۰
زمستان	۰	۰	۰	۱۰۰

بارش شدید باران و در نتیجه افزایش تجمع آب‌های سطحی و راکد می‌باشد. همچنین مطالعه حاضر نشان می‌دهد که بیش‌ترین موارد بیماری در فصل تابستان و اواخر فصل بهار به ثبت رسیده و کم‌ترین رخداد بیماری در فصل زمستان رخ داده است. از مهم‌ترین اهداف این تحقیق یافتن روابط میان بیماری

۵- نتیجه‌گیری و پیشنهادات

نتایج به‌دست آمده از بررسی‌های آماری، حاکی از وجود موارد وقوع بیماری لپتوسپیروز در نواحی مرکزی استان گیلان، مناطق شرقی استان مازندران و مناطق غربی استان گلستان، با آب و هوای معتدل و مرطوب،

خیلی قابل اتکا نباشند. نتایج این تحقیق می‌توانند ابزار مفیدی برای مدیران و تصمیم‌گیرندگان مسائل بهداشتی جهت برنامه‌ریزی در اولویت‌بندی‌های اختصاص بودجه، تصمیم‌گیری از طریق شناخت توزیع مکانی بیماری‌ها و عوامل موثر استخراج‌شده از مدل که بر توزیع آن‌ها تأثیرگذار بوده به منظور مبارزه با شیوع بیماری فراهم آورد. به عنوان کارهای آینده پیشنهاد می‌شود که از پارامترهای دیگر همچون کاربری زمین، نوع کشت، جهت شیب زمین، تعداد روزهای بارانی، طول دوره‌های گرما و غیره به منظور افزایش دقت و اعتمادپذیری نتایج حاصل از مدلسازی استفاده شود. همچنین پیشنهاد می‌شود که از سایر روش‌های مدلسازی همچون روش‌های رگرسیونی و شبکه‌های عصبی استفاده شود و نتایج حاصل از آنها با نتایج به‌دست آمده در این تحقیق مورد مقایسه قرار گیرد و مشخص شود که کدامیک از روش‌ها در مدلسازی بیماری لپتوسپیروز از قدرت و کارایی بیشتری برخوردار می‌باشند

تقدیر و تشکر

نویسندگان این مقاله از جناب آقای دکتر محمدرضا شیرزادی رئیس مرکز کنترل بیماری‌های واگیر وزارت بهداشت، درمان و آموزش پزشکی ایران بابت فراهم نمودن داده‌های بیماری لپتوسپیروز، نهایت تشکر و سپاس‌گذاری را دارند.

و عواملی که شرایط مساعدی را برای شیوع بیماری فراهم کرده‌اند، می‌باشد. از متغیرهای محیطی استفاده شده در مدلسازی چنین استنباط می‌شود که این متغیرها رابطه معنی‌داری با وقوع بیماری نشان می‌دهند. نتایج حاکی از آن بود که همبستگی مثبت و قوی بین متغیرهای رطوبت، دما و بارش و همبستگی منفی و ضعیف بین عناصر پوشش گیاهی و تعداد روزهای یخبندان با تعداد وقوع بیماری وجود دارد. همچنین نرخ شیوع بیماری در فصل گذشته ضعیف‌ترین همبستگی مثبت را با وقوع بیماری نشان می‌دهد. علاوه بر این صرف نظر از علامت‌های مثبت و منفی، متغیرهای تعداد روزهای یخبندان و نرخ شیوع بیماری در فصل گذشته ضعیف‌ترین همبستگی را با وقوع بیماری دارا می‌باشند. در این پژوهش توان مدلسازی الگوریتم یادگیری ماشینی SVM برای تولید نقشه آسیب‌پذیری بیماری لپتوسپیروز در سطح بخش‌های استان‌های شمالی ایران با استفاده از منحنی ROC مورد بررسی قرار گرفت. نتایج گویای این حقیقت بودند که مدل SVM عملکرد بسیار عالی در تهیه نقشه آسیب‌پذیری بیماری لپتوسپیروز داشته است. مهم‌ترین محدودیت این تحقیق مربوط به نواقص سیستم مراقبت و پایش بیماری در ایران است که چندان دقیق نیست و سالانه شمار قابل توجهی از موارد بیماری گزارش نمی‌شوند. محدودیت دیگر مربوط به دوره کوتاه (۶ سال) به دلیل فقدان داده‌های بیماری در سال‌های قبل از آن است که ممکن است نتایج

مراجع

- [1] F. Babamahmoodi, A. G. Âjemi, M. kalhor, A. R. Khalilian, and G. R. shfiei, "A sero epidemiological study of helicobacter pylori infection in Sari in 2001-2002," Journal of Mazandaran University of Medical Sciences, vol. 14, pp. 39-50, 2004.
- [2] M. C. Vega-Corredor and J. Opadeyi, "Hydrology and public health: linking human leptospirosis and local hydrological dynamics in Trinidad, West Indies," Earth Perspectives, vol. 1, p. 1, 2014.
- [3] F. Babamahmodi, N. Motamed, M. Mahdavi, F. Nickhah, and K. Qavi Bonyeh, "Seroepidemiological study of leptospirosis in Ghaemshahr Mazandaran province-Iran, Sept-Oct 2004," Journal of Mazandaran University of Medical Sciences, vol. 16, pp. 51-56, 2006.
- [4] H. Honarmand, S. Eshraghi, M.

- Khorramzadeh, R. Hartskeerl, F. Ghanaei, G. Abdolhoupour, et al., "Distribution of Human Leptospirosis in Guilan Province, Northern Iran," *Iranian Journal of Public Health*, vol. 36, pp. 68-72, 2007.
- [5] A. Rafiei, A. H. Zadeh-Omran, F. Babamahmoodi, R. A. Navaei, R. Valadan, and I. Sari, "Review of Leptospirosis in Iran," *Journal of Mazandaran University of Medical Sciences (JMUMS)*, vol. 22, 2012.
- [6] A. Sh, F. Baba Mahmoudi, N. Najafi, R. Ghasemian, S. Teimouri, and L. Shahbaz Nezhad, "Clinical and paraclinical findings of Leptospirosis in Mazandaran, June-September 2004," *J Mazand Univ Med Sci*, vol. 53, pp. 85-78, 2006.
- [7] M. AhangarCani, M. Farnaghi, and M. R. Shirzadi, "Predictive Map of Spatio-Temporal Distribution of Leptospirosis Using Geographical Weighted Regression and Multilayer Perceptron Neural Network Methods," *Journal of Geomatics Science and Technology*, vol. 6, pp. 79-98, 2016.
- [8] P. A. Burrough, "Principles of geographical information systems for land resources assessment," 1986.
- [9] Z. Massoomy and M. S. Mesgari, "PREDICTION OF SKIN CANCER EPIDEMIOLOGY FOR DECISION MAKING USING GEOSTATISTICAL ANALYSES," presented at the ASPRS 2006 Annual Conference, Reno, Nevada, 2006.
- [10] [10] S. Keola, M. Tokunaga, N. K. Tripathi, and W. Wisa, "Spatial surveillance of epidemiological disease: a case study in Ayutthaya province, Thailand," *GIS@ development*, 2002.
- [11] S. L. McLafferty, "GIS and health care," *Annual review of public health*, vol. 24, pp. 25-42, 2003.
- [12] D. A. Henderson, "Principles and lessons from the smallpox eradication programme," *Bulletin of the World Health Organization*, vol. 65, p. 535, 1987.
- [13] M. Zare, P. Shamszadeh, and A. Najjari, "Providing the opportunity to use GIS in decision-making in the health sector management," *Hakim*, vol. 9, pp. 58-63, 2006.
- [14] C. Barcellos and P. C. Sabroza, "The place behind the case: leptospirosis risks and associated environmental conditions in a flood-related outbreak in Rio de Janeiro," *Cadernos de Saúde Pública*, vol. 17, pp. S59-S67, 2001.
- [15] M. Bavia, D. Carneiro, H. da Costa Gurgel, C. M. Filho, and M. R. Barbosa, "Remote sensing and geographic information systems and risk of American visceral leishmaniasis in Bahia, Brazil," *Parassitologia*, vol. 47, p. 165, 2005.
- [16] Z. Pezeshki, M. Tafazzoli-Shadpour, A. Mansourian, B. Eshrati, E. Omidi, and I. Nejadqoli, "Model of cholera dissemination using geographic information systems and fuzzy clustering means: Case study, Chabahar, Iran," *Public health*, vol. 126, pp. 881-887, 2012.
- [17] C. L. Lau, A. C. Clements, C. Skelly, A. J. Dobson, L. D. Smythe, and P. Weinstein, "Leptospirosis in American Samoa—estimating and mapping risk using environmental data," *PLoS Negl Trop Dis*, vol. 6, p. e1669, 2012.
- [18] K. Vahdat, I. Nabipour, M. Motamedi, S. M. Jafary, A. Ghajary, M. H. Zafarmand, et al., "A sero-epidemiological survey on leptospirosis in the livestock breeders during the outbreak of haemorrhagic fever in domestic animals of the Helleh River area in 2004," *ISMJ*, vol. 8, pp. 53-59, 2005.
- [19] S. M. Alavi and M. M. Khoshkho, "Seroprevalence study of leptospirosis among rice farmers in khuzestan province, South west iran, 2012," *Jundishapur journal of microbiology*, vol. 7, 2014.
- [20] R. Golsha, B. Khodabakhshi, and A. RAHNAMA, "Leptospirosis in Golestan province in Iran (Reports of twelve cases)," 2007.
- [21] H. Faraji, "Prevalence of leptospirosis in

- Mazandaran province, using microagglutination and indirect fluorescence immunoassay," *Iranian Journal of Biology*, vol. 22, pp. 608-618, 2009.
- [22] R. Esmaili, F. Alhani, A. Hesamzadeh, R. Alizadeh Navaei, and M. Parsaei, "A report of 102 patients with leptospirosis in Mazandaran province between 2003 and 2008," *J Mazand Univ Med Sci*, vol. 19, pp. 72-75, 2009.
- [23] H. Honarmand, G. F. MANSOUR, A. HEYDARZADEH, and M. ASMAR, "Isolation and Serotyping of endemic leptospires of eastern part of flat area of Guilan province, Iran," *JOURNAL OF GORGAN UNIVERSITY OF MEDICAL SCIENCES*, vol. 11, pp. 53-59, 2009.
- [24] A. Mohammadinia, A. Alimohammadi, and R. Habibi, "Assessment of environmental factors associated with rural endemics of Leptospirosis in Guilan Province, Iran," 2015.
- [25] A. M. Nia, A. Alimohammadi, R. Habibi, and M. Shirzadi, "Spatial and Statistical Analysis of Leptospirosis in Guilan Province, Iran," *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 40, p. 497, 2015.
- [26] M. Rajabi, A. Mansourian, P. Pilesjö, and A. Bazmani, "Environmental modelling of visceral leishmaniasis by susceptibility-mapping using neural networks: a case study in north-western Iran," *Geospatial health*, vol. 9, pp. 179-191, 2014.
- [27] M. R. Shirzadi, A. Mollalo, and M. R. Yaghoobi-Ershadi, "Dynamic relations between incidence of zoonotic cutaneous leishmaniasis and climatic factors in Golestan Province, Iran," *Journal of arthropod-borne diseases*, vol. 9, p. 148, 2015.
- [28] A. Field, *Discovering statistics using IBM SPSS statistics*: Sage, 2013.
- [29] E. Akbari, E. Mayvaneh, A. Entezari, and M. Nazari, "Survey of the Role of Bioclimatic Factors in the Outbreak of Cutaneous Leishmaniasis," *Iranian Journal of Epidemiology*, vol. 10, pp. 65-74, 2014.
- [30] J. Benesty, J. Chen, Y. Huang, and I. Cohen, "Pearson correlation coefficient," in *Noise reduction in speech processing*, ed: Springer, 2009, pp. 1-4.
- [31] M. KHAZAEI and H. IZADBAKHS, "COMBINATION OF DEA AND PCA FOR FULL RANKING OF DECISION MAKING UNITS," *journal of Industrial Management*, vol. 1, pp. 55-70, 2009.
- [32] W. Lu, W. Wang, X. Wang, Z. Xu, and A. Leung, "Using improved neural network model to analyze RSP, NOx and NO2 levels in urban air in Mong Kok, Hong Kong," *Environmental Monitoring and Assessment*, vol. 87, pp. 235-254, 2003.
- [33] C. J. Burges, "A tutorial on support vector machines for pattern recognition," *Data mining and knowledge discovery*, vol. 2, pp. 121-167, 1998.
- [34] Y. Wang, Z. Xue, G. Shen, and J. Xu, "PRINTR: prediction of RNA binding sites in proteins using SVM and profiles," *Amino acids*, vol. 35, pp. 295-302, 2008.
- [35] H. Yu and S. Kim, "Svm tutorial—classification, regression and ranking," in *Handbook of Natural computing*, ed: Springer, 2012, pp. 479-506.
- [36] B. Schölkopf, A. Smola, K.-R. Müller, C. Burges, and V. Vapnik, "Support vector methods in learning and feature extraction," 1998.
- [37] K.-B. Duan and S. S. Keerthi, "Which is the best multiclass SVM method? An empirical study," in *International Workshop on Multiple Classifier Systems*, 2005, pp. 278-285.
- [38] Y. Liu and Y. F. Zheng, "One-against-all multi-class SVM classification using reliability measures," in *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005.*, 2005, pp. 849-854.
- [39] G. Anthony, H. Gregg, and M. Tshilidzi,

"Image classification using SVMs: one-against-one vs one-against-all," arXiv preprint arXiv:0711.2914, 2007.

- [40] J. Milgram, M. Cheriet, and R. Sabourin, "'One Against One' or 'One Against All': Which One is Better for Handwriting Recognition with SVMs?," in Tenth International Workshop on Frontiers in Handwriting Recognition, 2006.
- [41] N. A. Obuchowski, "Receiver Operating Characteristic Curves and Their Use in Radiology 1," *Radiology*, vol. 229, pp. 3-8, 2003.
- [42] S. J. Phillips, R. P. Anderson, and R. E. Schapire, "Maximum entropy modeling of species geographic distributions," *Ecological modelling*, vol. 190, pp. 231-259, 2006.
- [43] D. M. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," 2011.



Exploring the Impact of Topographical and Climate Factors on Generation of the Vulnerability-map of Leptospirosis

Mehrdad Ahangarani¹, Mahdi Farnaghi^{*2}, Mohammad Reza Shirzadi³

1- M.Sc student of geographic information system, Faculty of Geodesy and Geomatics, K. N. Toosi University of Technology

2- Assistant Professor, Faculty of Geodesy and Geomatics, K. N. Toosi University of Technology

3- Center for Disease Control (CDC), Ministry of Health and Medical Education

Abstract

Leptospirosis is one of the most widespread zoonotic disease caused by *Leptospira* bacteria. It is found wherever human is in direct or indirect contact with *Leptospira* bacteria thorough infected animals as well as contaminated soil or water. The disease is mostly found in tropical, subtropical, hot, and humid areas. The main objectives of this study are to investigate the seasonality relations between the topographical and climate factors, including altitude, slope, vegetation, average temperature, average humidity, precipitation and number of freezing days and incidence of Leptospirosis as well as modelling of Leptospirosis using support vector machine at the district level in Northern provinces of Iran. Pearson's correlation analysis was conducted to examine the type and strength of relationships between the topographical and climate variables and Leptospirosis incidence. Results of Pearson's correlation analysis indicate that average humidity, average temperature and rainfall were the most influential environmental factors which as effect on prevalence of Leptospirosis in the study area. Statistical analysis showed that most cases of the Leptospirosis prevalence have been recorded in the late spring and summer. On the other hand, the lowest incidences have occurred in winter. Also, high distribution of leptospirosis mainly located in the central areas of Guilan province, the eastern parts of Mazandaran province and western regions of Golestan province with a mild and humid climate and abundant rainfall. Eventually, performance of support vector machine (SVM) model evaluated by area under the ROC curve. The output maps showed that SVM model has excellent performance in the vulnerability mapping of Leptospirosis.

Key words: Pearson's analysis, geographic information system, support vector machine, ROC curve, Leptospirosis

Correspondence Address Geospatial Information System Group, Faculty of Geodesy and Geomatics Engineering, K.N.Toosi University of Technology, Tehran, Iran
Tel : +98 912 243 2847
Email: farnaghi@kntu.ac.ir