

پیش‌بینی مکانی-زمانی مناطق پرخطر بیماری کرونا با استفاده از روش‌های رگرسیون وزندار جغرافیایی و شبکه عصبی پرسپترون چندلایه

میلاذ ایلخانی خسروشاهی^۱، محمد کریمی^{۲*}، ندا کفاش چرندابی^۳

^۱ دانشجوی کارشناسی ارشد سیستم‌های اطلاعات مکانی - دانشکده مهندسی نقشه‌برداری -

دانشگاه صنعتی خواجه نصیرالدین طوسی

milkhanikh@email.kntu.ac.ir

^۲ دانشیار دانشکده مهندسی نقشه‌برداری - دانشگاه صنعتی خواجه نصیرالدین طوسی

mkarimi@kntu.ac.ir

^۳ استادیار گروه نقشه‌برداری - دانشکده فنی مهندسی موند - دانشگاه تبریز

n_kaffash@tabrizu.ac.ir

(تاریخ دریافت: دی‌ماه ۱۴۰۰، تاریخ تصویب: دی‌ماه ۱۴۰۱)

چکیده

یکی از مسری‌ترین بیماری‌های سال‌های اخیر کووید ۱۹ (کرونا) است که از اواخر سال ۲۰۱۹ میلادی از شهر ووهان - چین به سراسر جهان شیوع پیدا کرد و جهان را با بحران‌های بسیار مواجه و تاثیر عمیقی در دنیا و زندگی روزمره انسان‌ها گذاشت. در بیشتر افراد آلوده به این بیماری باعث بروز علائم تنفسی شده که شدت آن بستگی به سیستم ایمنی بدن فرد دارد. اهداف اصلی این تحقیق، کشف خوشه‌ها و پیش‌بینی مناطق پرخطر بیماری کرونا، مقایسه کارایی دو روش پیشنهادی و تعیین پارامترهای موثر به تفکیک شهرستان می‌باشد. در این تحقیق به منظور بررسی الگوی توزیع نرخ بروز بیماری و خوشه‌ها به ترتیب از شاخص موران و شاخص تحلیل نقطه داغ و به منظور تعیین پارامترهای موثر بیماری از ضریب همبستگی پیرسون استفاده شد. در این تحقیق داده‌های آماری بیماری کرونا استان آذربایجان شرقی در سطح شهرستان به همراه داده‌های محیطی و توپوگرافی، سلامت و بهداشت، اقتصادی و تسهیلات شهری در محدوده زمانی ۳ اسفند ۱۳۹۸ تا ۳۰ آبان ۱۳۹۹، بصورت هفتگی جمع‌آوری گردید. مطابق نتایج بدست‌آمده، روند بروز بیماری کرونا طی این بازه زمانی دو پیک را پشت سر گذاشته و براساس نقشه‌های بدست‌آمده از دو مدل در برخی هفته‌ها مدل GWR و در برخی مدل MLP برتری بود؛ هم‌چنین برای مدل GWR مقدار شاخص خوبی برازش ۰/۸۹۸۵ و خطای جذر میانگین مربعات نرمال شده ۰/۰۸۲۲ و برای مدل MLP نیز بترتیب برابر با ۰/۸۲۲۶ و ۰/۱۳۴۰ بدست آمد که نشان می‌دهد روش GWR مناسب‌تر است. آنالیز تحلیل حساسیت پارامترهای مختلف نشان داد که پارامترهای نرخ بروز کرونا هفته ماقبل و سرعت باد مهم‌تر از سایر پارامترهای مدلسازی شده در این مسئله می‌باشد. در این تحقیق پارامترهای موثر به تفکیک هر شهرستان استخراج گردید و یک مدل محلی ارائه شد که در مقایسه با حالت کلی، مدل محلی دقت بهتری نسبت به مدل کلی روش MLP دارا بود.

واژگان کلیدی: کووید ۱۹ (کرونا و بروس)، توزیع مکانی زمانی، پیش‌بینی، شبکه عصبی پرسپترون چند لایه، رگرسیون وزندار جغرافیایی، سیستم اطلاعات مکانی.

* نویسنده رابط

۱- مقدمه

در اواخر دسامبر سال ۲۰۱۹، ویروس جدید معروف به کرونا ویروس سندرم حاد تنفسی^۱ یافت شد که سرمنشاء آن شهر ووهان در کشور چین بود. تا ۱۱ ژوئن سال ۲۰۲۰، کووید ۱۹ به سرعت در اکثر کشورها در سراسر جهان گسترش یافته و ۷,۵۳۹,۴۰۲ نفر را تحت تأثیر قرار داده است. نکته اصلی حمله کووید ۱۹ آسیب به سیستم تنفسی و حفره‌های کوچک موجود در آن است. ویروس وارد خون می‌شود، پس از عفونت‌ریه، در کلیه جمع می‌شود و می‌تواند به سلول‌های کلیوی ساکن آسیب وارد کند. دوره نقاهت ویروس کووید ۱۹ تقریباً ۳-۱۴ روز است [۱].

به منظور پیش‌بینی و کنترل بیماری‌های همه‌گیر، مدل‌های محاسباتی بسیاری توسط محققین ارائه شده است. افزایش دسترسی به منابع کامپیوتری، مدلسازی گسترش همه‌گیری را در چندین سطوح فردی، ناحیه ای و کشوری تسهیل نموده است. مدل‌های کامپیوتری باعث بهبود نمایش، فهم ساختارهای پیچیده اجتماعی و الگوهای ناهمگون درون شبکه‌های ارتباطی بین افراد و تخمین پویایی انتقال بیماری می‌شوند [۲]. با توجه به شدت بالینی کووید ۱۹، میزان شیوع آن و نگرانی عمومی، تعیین محدودیت‌های دقیق اپیدمیولوژیک و مدلسازی آن‌ها ضروری است [۳].

یکی از مهمترین خواص اپیدمی‌ها گسترش مکانی آن‌ها است. یک ویژگی که عمدتاً به مکانیسم اپیدمی، نیروی انسانی و استراتژی کنترل بستگی دارد که می‌توان از سامانه اطلاعات مکانی^۲ و آمار مکانی برای پاسخ به آن‌ها استفاده نمود [۴]. پیش‌بینی شیوع مکانی و روند زمانی شیوع بیماری و شناسایی مناطق جغرافیایی و گروه‌های آسیب‌پذیر برای خدمات درمانی ضروری است و به تصمیم‌گیرندگان کمک می‌کنند تا با برنامه‌ریزی‌های بهتر، سیاست‌های بهداشتی را بهتر دنبال کنند [۵].

در مدلسازی و پیش‌بینی مکانی زمانی مناطق پرخطر بیماری کرونا مطالعات مختلفی انجام گرفته است که در ادامه به اختصار ذکر می‌شوند. تحقیقات آقای گوان و همکاران در

ژانویه سال ۲۰۲۰ برای اولین بار بیانگر اهمیت استفاده از GIS برای تحلیل مکانی کووید ۱۹ بود. آن‌ها داده‌های مربوط به ۱۰۹۹ بیمار را در چین جمع‌آوری و تحلیل نمودند. آقای چن و همکاران در سال ۲۰۲۰ میلادی با استفاده از یک مدلسازی مکانی زمانی بیزین^۳، توزیع مکانی کووید ۱۹ و ارتباط آن‌ها با خروج مردم ووهان در مراحل اولیه بیماری همه‌گیر را تعیین کردند [۴]. آقای وانگ و همکاران در سال ۲۰۲۰ دریافتند که پارامترهای درجه حرارت و رطوبت نسبی بالا به نظر می‌رسد که انتقال ویروس را کاهش دهد [۴].

آقای پورقاسمی و همکاران در سال ۱۳۹۹ روند انتشار بیماری کرونا در ایران را با استفاده از روش یادگیری ماشینی^۴ (MLT و جنگل تصادفی)^۵ (RF) پیش‌بینی و با روند جهانی موجود مقایسه نمودند. در این تحقیق عنوان شد که تراکم جمعیت استان‌ها و مسیر مسافرت مهمترین عواملی هستند که باعث شیوع ویروس می‌شوند [۲]. ایشان در تحقیقی دیگر، در سال ۲۰۲۰، از الگوریتم SVM^۶ و مدل چندجمله‌ای ARIMA^۷ جهت بررسی رفتار همه‌گیری بیماری در استان فارس و ایران استفاده و نشان دادند که فاصله از ایستگاه‌های اتوبوس، بیمارستان‌ها و نانوایی‌ها دارای اهمیت بالاتری هستند، در حالی که عوامل اقلیمی کمترین تأثیر را داشتند و مدل تولید شده با استفاده از SVM دارای دقت بهتری بود [۶].

رامیرز و لی در سال ۲۰۲۰ با هدف بررسی الگوهای مکانی بروز کووید ۱۹ در سطح مناطق جمعیتی ایالت کلرادو در کشور آمریکا با محدودیت زمانی ۵ روزه، از الگوریتم IDW استفاده نمودند. نتایج این تحقیق نشان داد که میزان شیوع کرونا با عوامل تراکم جمعیت، درآمد سرانه، اقلیت، ساختارهای چند واحدی و میزان بیماران آسم رابطه مثبت و با عامل خانه‌های سیار رابطه منفی داشت [۷]. آقای ملین و همکاران در سال ۲۰۲۰ در یک تحقیق تکامل مکانی همه‌گیری ویروس کرونا در سراسر جهان را با استفاده از نوع خاصی از شبکه عصبی نظارت‌نشده تحلیل کردند [۸].

آقای ملالو و همکاران در سال ۲۰۲۰ با هدف مدلسازی مکانی میزان شیوع کرونا در کشور آمریکا، تاثیر ۳۵ پارامتر

۳ Bayesian

۴ Machine learning technique

۵ Random forest

۶ Support vector mechine

۷ Autoregressive Integrated Moving Average

۱ Severe Acute Respiratory Syndrome Coronavirus

۲ Geographic Information System (GIS)

گذارند و مدل MGWR بهترین عملکرد را داشته و می‌توان با تهیه نقشه و شناسایی مردم آسیب‌پذیری که در معرض خطر بالای این بیماری هستند کمک کند [۲۰].

اهمیت منطقه مورد مطالعه در بررسی بیماری کرونا و وابستگی عوامل بیماریزا به شرایط محیطی و اقلیمی حاکم بر آن منطقه، میزان تاثیر عوامل مورد بررسی در بروز بیماری در منطقه مورد مطالعه و پیش‌بینی آینده از ضروریات این تحقیق بشمار می‌رود. با توجه به اینکه عوامل موثر بر بیماری کووید ۱۹ و نحوه گسترش این بیماری در کشورها و ایالت‌های مختلف متفاوت بوده است، این تحقیق، با هدف کشف خوشه‌های مکانی زمانی بیماری کووید ۱۹ و همچنین تهیه نقشه پیش‌بینی توزیع مکانی زمانی مناطق پرخطر این بیماری و مقایسه کارایی دو روش پیشنهادی شبکه عصبی پرسپترون چندلایه و رگرسیون وزندار جغرافیایی برای یک استان نمونه در کشور ایران، استان آذربایجان شرقی، تعریف شد. همچنین در این مطالعه قصد داریم علاوه بر تعیین کلی پارامترهای موثر، میزان موثر بودن پارامترها به تفکیک شهرستان را بدست آورده و یک مدل محلی ارائه دهیم. در این تحقیق به منظور بررسی الگوی توزیع نرخ بروز بیماری کرونا و خوشه‌های مکانی این بیماری از شاخص موران عمومی و شاخص تحلیل نقطه داغ و به منظور تعیین پارامترهای موثر بیماری از ضریب همبستگی پیرسون استفاده شد.

در ادامه بعد از تشریح مفاهیم زمین آمار و معرفی روش‌های شبکه عصبی پرسپترون چندلایه و رگرسیون وزندار جغرافیایی، نحوه پیش‌بینی مکانی زمانی مناطق پرخطر بیماری کرونا در استان آذربایجان شرقی بررسی و تحلیل شده‌است.

۲- مواد و روش‌ها

در این فصل بطور اجمالی به مبانی نظری تحقیق پرداخته و به معرفی روش‌های تحلیل مکانی و روش‌های مدل‌سازی بیماری کرونا پرداخته می‌شود.

۲-۱- آمار مکانی

آمار مکانی شامل مجموعه‌ای از تکنیک‌ها برای توصیف و مدل‌سازی داده‌های مکانی است. این تکنیک‌ها در واقع از ریاضیات خود از فضا - مساحت، طول، مجاورت، جهت‌گیری یا روابط مکانی استفاده می‌کنند [۲۱]. زمین‌آمار ابزاری را

برای آنالیز مشترک مولفه‌های مکان و زمان فراهم می‌کند. مدل‌های مکانی - زمانی زمین آماری یک چارچوب احتمالی برای تجزیه و تحلیل داده‌ها و پیش‌بینی‌ها ایجاد می‌کند که بر وابستگی مکانی و زمانی مشترک بین مشاهدات بنا می‌شود [۲۲]. تعیین الگوی فضایی بیماری‌ها با استفاده از تهیه نقشه بیماری‌ها یک مشکل عمده در اپیدمیولوژی است. به طور کلی، به دلیل شباهت در سبک زندگی، شرایط محیطی، عوامل ژنتیکی و غیره، انتظار می‌رود الگوهای مشابهی برای بروز بیماری در مناطق همسایه وجود داشته باشد [۲۳]. در این مطالعه به منظور بررسی الگوی توزیع بیماری در سطح منطقه مورد مطالعه از شاخص موران عمومی و جهت کشف خوشه‌های مکانی بیماری از تحلیل نقطه داغ استفاده گردید که در ادامه به تشریح هر یک می‌پردازیم.

۲-۱-۱- شاخص موران عمومی

به کمک شاخص موران عمومی می‌توان میزان پراکندگی یا جمع‌شدگی داده‌های مکانی را در فضا پایش کرد. برای این شاخص فرض صفر بیان می‌کند که پدیده مورد تحلیل به طور تصادفی در میان عارضه‌های منطقه مورد مطالعه شما توزیع می‌شود. زمانی که مقدار احتمال برگردانده شده توسط این شاخص از نظر آماری معنی‌دار باشد، می‌توان فرض صفر را رد نمود. مقدار این شاخص با استفاده از رابطه (۱) محاسبه می‌گردد [۲۴]:

$$I = \frac{N}{\sum_{i=1}^N \sum_{j=1}^N w_{ij}} * \frac{\sum_{i=1}^N \sum_{j=1}^N w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^N (x_i - \bar{x})^2} \quad (1)$$

که x_i, x_j بترتیب بیانگر تعداد موارد بیماری در i, j امین مقیاس مکانی (شهرستان) هستند. \bar{x} نشانگر میانگین تعداد مبتلایان هر شهرستان، N تعداد شهرستان‌ها در منطقه مورد مطالعه و w_{ij} نشان‌دهنده وزن مکانی بین موارد i, j می‌باشند [۲۴]. مقدار شاخص موران بین ۱- و ۱+ بوده بطوری که بالای صفر نمایانگر خودهمبستگی مثبت و مقادیر زیر صفر نمایانگر خودهمبستگی منفی بوده و در صورتی که برابر صفر باشد نمایانگر عدم وجود خودهمبستگی مکانی (توزیع داده تصادفی) می‌باشد. هرچه مقدار شاخص موران به عدد ۱ نزدیک تر باشد توزیع داده‌ها دارای الگوی خوشه‌ای و اگر به مقدار ۱- نزدیک تر باشند توزیع داده‌ها الگوی یکنواخت دارند (معیار وزندهی همسایگی دو چندضلعی می‌باشد).

یک فاکتور نرمالیزاسیون عمل می‌کند) تعریف می‌شود و می‌توان آن را به صورت رابطه (۵) تعریف کرد [۲۶].

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (5)$$

که \bar{x} میانگین متغیر x و \bar{y} میانگین متغیر y بوده و از روابط (۳) و (۴) بدست می‌آیند. ضریب همبستگی پیرسون در محدوده ۱- تا ۱ بوده و در تبدیلات خطی هر دو متغیر ثابت نیستند. ضریب همبستگی پیرسون قدرت رابطه خطی بین دو متغیر تصادفی x و y را نشان می‌دهد. علامت ضریب همبستگی در صورتیکه متغیرها مستقیماً با هم مرتبط باشند مثبت است و در صورتیکه رابطه عکس با هم داشته باشند، منفی است. اگر این ضریب برابر صفر باشد، گفته می‌شود که x و y با هم ارتباط ندارند. هر چه این مقدار به ۱ نزدیکتر باشد، نزدیکتر بودن اندازه گیری‌ها به یک رابطه خطی قوی‌تر است [۲۶].

۲-۳- رگرسیون وزندار جغرافیایی

رگرسیون وزندار جغرافیایی (GWR) توسط براندسون، فودرینگهام و کارلتون (۱۹۹۶) به رشته جغرافیا معرفی شد تا به بررسی پتانسیل روابط در یک مدل رگرسیون برای تغییر در فضای جغرافیایی، یا آنچه اصطلاحاً ناپستیایی پارامتری نامیده می‌شود بپردازد. GWR مبتنی بر روش غیرپارامتری رگرسیون وزندار محلی است که در آمار برای کاربردهای برازش منحنی و هموارسازی ساخته شده‌است، که در آن پارامترهای رگرسیون محلی با استفاده از زیرمجموعه‌ای از داده‌های تقریبی به یک نقطه تخمینی مدل در فضای متغیر برآورد می‌شوند. در GWR به جای فضای متغیر از زیرمجموعه‌ای از داده‌های نزدیک به مکان کالیبراسیون مدل در فضای جغرافیایی استفاده می‌کند [۲۷].

این روش روابط بین متغیرهایی که با عوارض جغرافیایی مربوط می‌شوند را مدل‌سازی نموده و به ما امکان پیش‌بینی مقادیر متغیرهای نامعلوم و فهم بهتر عواملی که بروی یک متغیر تاثیر می‌گذارند را می‌دهد. روش های رگرسیون به ما امکان می‌دهند که روابط بین متغیرها را بررسی نموده و میزان شدت و ضعف آن روابط را اندازه‌گیری نماییم. این روش این کار را از طریق تهیه معادلات رگرسیون جداگانه برای هر عارضه با ملاحظه متغیرهای وابسته و مستقلی که در طول باند (محدوده) عارضه قرار می‌گیرند انجام می‌دهد.

۲-۱-۲- شاخص Getis Ord G_i^* (Hotspot Analysis)

شاخص G_i^* Getis-Ord یا تحلیل نقطه داغ الگوهای مختلف خوشه‌بندی مکانی مانند نقاط داغ، پرخطر و نقاط سرد را در کل منطقه مورد مطالعه با اهمیت آماری مشخص می‌کند. در این تحلیل بررسی در یک همسایگی صورت می‌گیرد و از آنجا که دنبال یک خوشه بوده نمی‌تواند فاصله‌های دور را بررسی کند. مقادیر بالا/پایین زمانی اهمیت پیدا می‌کند که همسایه‌ها هم در همان محدوده باشند. برای هر همسایگی جهت مقایسه ابتدا وزن پیدا می‌کند تا اختلاف آن‌ها را نسبت به همدیگر مقایسه کند اگر این اختلاف کم باشد در می‌یابیم که نقاط داغ یا سرد همگی کنار هم هستند. مقدار G_i^* از رابطه (۲) بدست می‌آید [۲۵]:

$$G_i^* = \frac{\sum_{j=1}^n w_{ij}x_j - \bar{X} \sum_{j=1}^n w_{ij}}{S \sqrt{\frac{[n \sum_{j=1}^n w_{ij}^2 - (\sum_{j=1}^n w_{ij})^2]}{n-1}}} \quad (2)$$

که x_j مقدار داده توصیفی مربوط به عارضه $w_{i,j}$ وزن مکانی بین عارضه i و j بوده، n برابر با تعداد کل عارضه‌ها می‌باشد و S, \bar{X} از روابط (۳) و (۴) بدست می‌آیند:

$$\bar{X} = \frac{\sum_{j=1}^n x_j}{n} \quad (3)$$

$$S = \sqrt{\frac{\sum_{j=1}^n x_j^2}{n} - (\bar{X})^2} \quad (4)$$

برای کاربردهای مربوط به آنالیز G_i^* می‌توان به آنالیز جرم، اپیدمیولوژی، آنالیز الگوی رأی‌گیری، جغرافیای اقتصادی، تجزیه و تحلیل خرده فروشی، تجزیه و تحلیل حوادث ترافیکی و جمعیت‌شناسی یافت [۲۵].

۲-۲- همبستگی پیرسون

ضریب همبستگی پیرسون یک اندازه‌گیری از وابستگی خطی بین دو متغیر تصادفی است. از نظر تاریخی، اولین معیار رسمی از همبستگی بوده و هنوز هم یکی از پرکاربردترین معیارهای رابطه‌ای است. ضریب همبستگی پیرسون دو متغیر X و Y به طور رسمی به عنوان کوواریانس دو متغیر تقسیم بر حاصل انحراف معیار آن‌ها (که به عنوان

شکل و اندازه این محدوده بستگی به ورودی‌هایی دارد که کاربر تعیین می‌نماید و معمولاً یکی از روش‌های کرنل، روش طول باند^۱، فاصله و تعداد عوارض می‌باشند [۲۸]. باتوجه به اینکه رابطه ناهمگنی بین بیماری و پارامترهای موثر وجود دارد، برای در نظر گرفتن ناهمگنی باید از برخی روش‌ها استفاده شود. رگرسیون وزندار جغرافیایی یک رویکرد رایج است که می‌تواند با در نظر گرفتن تنوع ضرایب در مکان‌های متنوع در سراسر منطقه مورد مطالعه، ناهمگنی را حل کند. یکی از مزایای GWR در نظر گرفتن مکان پارامترها به عنوان ورودی برای بهبود قابلیت پیش‌بینی مکانی و کاهش اثر ناهمگنی است. GWR یک رویکرد کارآمد برای مدل‌سازی در زمینه‌های مختلف مطالعه، به ویژه مدل‌سازی و پیش‌بینی بیماری است [۲۹]. رگرسیون وزندار جغرافیایی از رابطه (۶) بدست می‌آید:

$$y_i = \beta_{i0} + \sum_{j=1}^m \beta_{ij} X_{ij} + \varepsilon_i, i = 1, 2, \dots, n \quad (6)$$

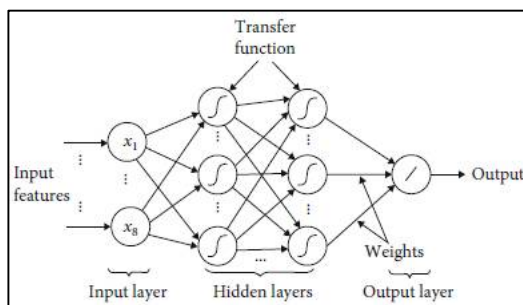
که در شهرستان i ، y_i برای نرخ شیوع کووید ۱۹ مقدار دارد، داشته، β_{ij}, β_{i0} ضرایب رگرسیون بوده، مقدار پارامتر توضیحی ε_i یک اصطلاح خطای تصادفی است [۹].

۲-۴- شبکه عصبی پرسپترون چندلایه^۲

شبکه عصبی که تحت عنوان شبکه عصبی مصنوعی شناخته می‌شود، نوعی الگوریتم یادگیری ماشین است که از مغز بیولوژیکی الهام گرفته شده است. شبکه عصبی مصنوعی یک مدل غیرخطی است که بر تعیین وابستگی بین پارامترهای ورودی و خروجی با شبیه‌سازی واحدهای پردازشی کاملاً متصل (نورون‌ها) سیستم عصبی انسان متمرکز است. شبکه‌های عصبی شامل سه لایه ورودی، پنهان و خروجی هستند و از اتصالات وزنی بین ورودی و خروجی تشکیل شده‌اند [۲۹].

شبکه عصبی پرسپترون چند لایه (MLP) یک کلاس مهم ANN^۳ است. این یک مدل ANN پیشخوان مبتنی بر آموزش نظارت‌شده است. یک MLP از مجموعه‌ای از

واحدهای ورودی (لایه ورودی)، یک یا چند مجموعه از گره‌های محاسباتی (لایه های پنهان) و یک مجموعه از گره‌های محاسباتی/خروجی (لایه خروجی) تشکیل شده است. در یک شبکه عصبی پرسپترون چندلایه، هر نورون در هر لایه به تمامی نورون‌های لایه قبل متصل می‌باشد. یک نورون روی لایه i روی یک MLP پیشروی^۴ دقیق همیشه به یک نورون روی لایه $i+1$ متصل می‌شود ضمناً تابع فعال‌سازی وجود دارد که خروجی یک نورون را از نظر ترکیبی خطی از ورودی‌ها تعریف می‌کند بعبارتی از تابع فعال‌ساز برای محدود کردن خروجی نورون استفاده می‌شود (شکل ۱). به منظور آموزش شبکه‌های عصبی پرسپترون چندلایه با هدف تنظیم پارامترهای شبکه و نزدیکی خروجی شبکه به خروجی واقعی، معمولاً از قانون یادگیری پس انتشار خطا^۵ استفاده می‌شود، که از دو مرحله تشکیل شده است: در ورودی پیش‌رو، یک بردار ورودی به شبکه ارائه می‌شود و به سمت خروجی انتشار می‌یابد. در مرحله پس انتشار، خروجی شبکه با خروجی مورد نظر مقایسه می‌شود. سپس، وزن شبکه مطابق با قانون تصحیح خطا تنظیم می‌شود، از این رو به آن 'پس‌انتشار' گفته می‌شود [۳۰].



شکل ۱- ساختار کلی شبکه عصبی پرسپترون چندلایه [۱۶]

۳- پیاده‌سازی و نتایج

منطقه مورد مطالعه در این پژوهش، استان آذربایجان شرقی بوده که یکی از استان‌های مهم کشور ایران بوده و سهم قابل توجهی در آمار مبتلایان کشور داشته که در محدوده‌ی شمال غرب کشور جای گرفته است (شکل ۲). این استان دارای ۴۵۴۹۱ کیلومتر مربع مساحت می‌باشد که

۴ feed forward
۵ Back Propagation

۱ Bandwidth Method
۲ Multi Layer Perceptron (MLP)
۳ Artificial neural network

که برای این ۶ هفته مطابق موارد ذکر شده در مورد مقدار شاخص موران و همچنین مقادیر p-value که نمایانگر رد شدن و یا نشدن فرض صفر می‌باشد و هر چه کوچکتر و به صفر نزدیکتر، فرض صفر رد خواهد شد و در غیر اینصورت فرض صفر رد نخواهد شد. مقادیر z-score نیز نشانگر اختلاف مقادیر واقعی از میانگین بوده و هر چه این مقدار بزرگتر باشد خوشه‌ای بودن شدیدتر احساس خواهد شد.

توزیع بیماری کرونا از شاخص عمومی موران برای هر هفته بصورت مجزا استفاده گردید که نتایج مربوطه در جدول (۳) نشان داده شده‌است. خروجی این شاخص نوع توزیع (خوشه‌ای، یکنواخت یا تصادفی) را نشان می‌دهد. مطابق نتایج بدست‌آمده تنها برای ۶ هفته (هفته ۵۱ سال ۹۸، هفته-های ۱۹، ۱۸، ۱۷، ۴، ۵ سال ۹۹) الگوی توزیع نرخ بروز بیماری خوشه‌ای بود و در مابقی هفته‌ها الگوی توزیع خوشه‌ای نبوده

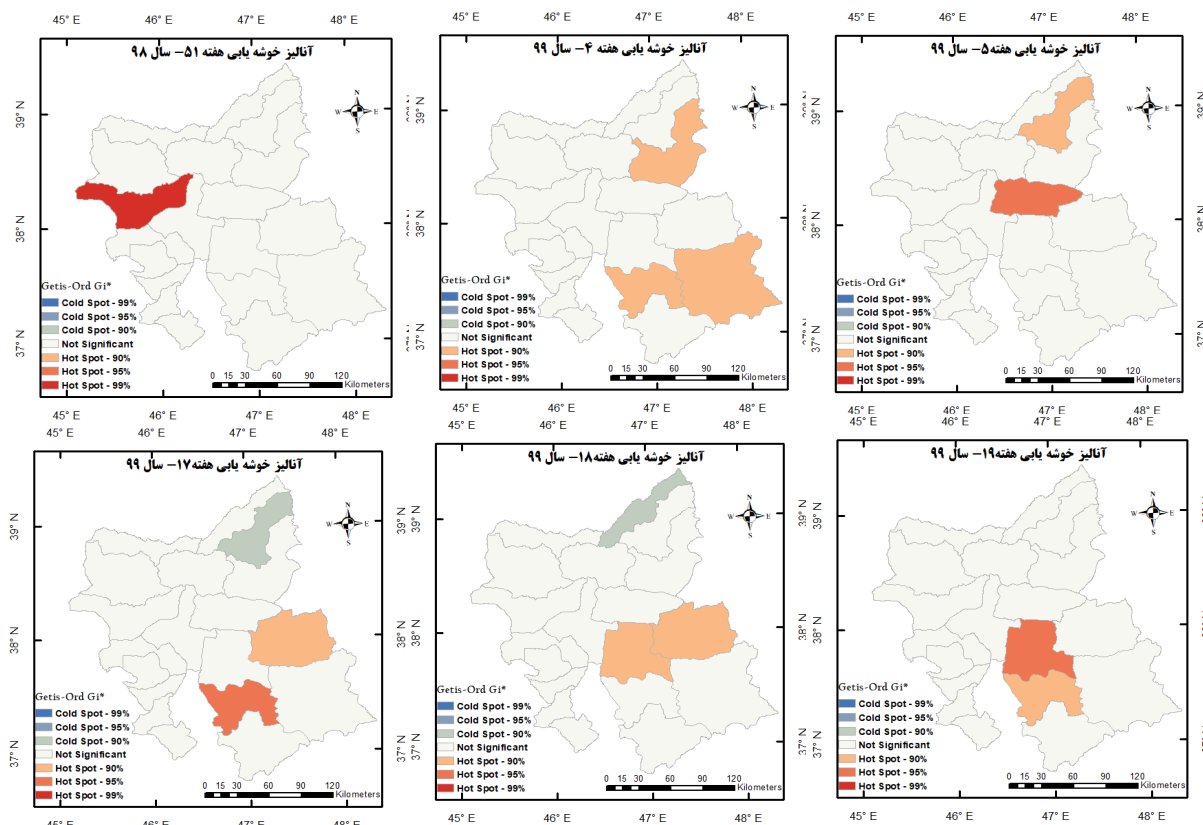
جدول ۳- نتایج آزمون موران عمومی توزیع نرخ بروز هفتگی کرونا

هفته	شاخص موران	p-value	z-score	توزیع	هفته	شاخص موران	p-value	z-score	توزیع
۴۹	-۰/۰۲۱۱	۰/۶۳۱۰	۰/۴۸۰۲	تصادفی	۱۶	۰/۱۱۲۶	۰/۳۵۴۷	۰/۹۲۵۳	تصادفی
۵۰	-۰/۲۳۹۶	۰/۲۷۳۶	-۱/۰۹۴۶	تصادفی	۱۷	۰/۳۱۴۷	۰/۰۴۷۱	۱/۹۸۴۷	خوشه‌ای
۵۱	-۰/۲۳۰۶	۰/۰۸۸	۱/۷۰۲۱	خوشه‌ای	۱۸	۰/۴۳۱۴	۰/۰۰۹۷	۲/۵۸۲۹	خوشه‌ای
۵۲	-۰/۱۵۸۳	۰/۲۰۹۳	۱/۲۵۵۴	تصادفی	۱۹	۰/۵۰۲۴	۰/۰۰۲۵	۳/۰۱۵۶	خوشه‌ای
۵۳	-۰/۱۴۱۸	۰/۶۴۹۰	-۰/۴۵۵۰	تصادفی	۲۰	-۰/۱۳۰۱	۰/۶۸۹۲	-۰/۳۹۹	تصادفی
۱	-۰/۱۲۵۷	۰/۳۴۰۵	۰/۹۵۳۰	تصادفی	۲۱	۰/۲۰۱	۰/۱۷۲۶	۱/۳۶۳۸	تصادفی
۲	-۰/۰۷۲۷	۰/۹۲۸۰	-۰/۰۹۰۳	تصادفی	۲۲	-۰/۱۳۴۲	۰/۶۷۲۶	-۰/۴۲۲۴	تصادفی
۳	-۰/۰۲۲۱	۰/۸۴۲۵	۰/۱۹۸۶	تصادفی	۲۳	۰/۰۱۴۹	۰/۷۰۹۹	۰/۳۷۱۸	تصادفی
۴	-۰/۲۶۷۵	۰/۰۸۵۹	۱/۷۱۷۱	خوشه‌ای	۲۴	-۰/۱۵۴۳	۰/۵۸۰۷	-۰/۵۵۲۲	تصادفی
۵	-۰/۳۲۵۰	۰/۰۴۳۰	۲/۰۲۳۲	خوشه‌ای	۲۵	-۰/۲۵۸۳	۰/۲۶۹۸	-۱/۱۰۳۳	تصادفی
۶	-۰/۲۱۴۰	۰/۱۵۱۱	۱/۴۳۵۶	تصادفی	۲۶	۰/۰۰۰۴	۰/۷۴۶۸	۰/۳۲۲۷	تصادفی
۷	-۰/۳۱۱۲	۰/۱۶۶۷۹	-۱/۳۸۲۵	تصادفی	۲۷	-۰/۱۱۰۴	-۰/۳۰۵۴	۰/۷۶۰۰	تصادفی
۸	-۰/۱۹۱۷	۰/۴۲۰۷	-۰/۸۰۵۰	تصادفی	۲۸	۰/۰۷۸۲	۰/۴۵۳۵	-۰/۷۴۹۵	تصادفی
۹	-۰/۳۱۰۵	۰/۱۵۳۵	-۱/۴۲۶۹	تصادفی	۲۹	۰/۰۲۲۵	۰/۶۷۲۴	-۰/۴۲۲۷	تصادفی
۱۰	-۰/۰۳۰۷	۰/۸۸۱۳	۰/۱۴۹۲	تصادفی	۳۰	-۰/۱۱۴۱	۰/۷۵۱۷	-۰/۳۱۶۳	تصادفی
۱۱	-۰/۰۸۶۴	۰/۸۳۴۲	-۰/۲۰۹۲	تصادفی	۳۱	-۰/۳۴۲۰	۰/۱۱۵۰	-۱/۵۷۵۹	تصادفی
۱۲	-۰/۰۶۶۲	۰/۹۵۱۶	-۰/۰۶۰۵	تصادفی	۳۲	-۰/۲۶۳۲	۰/۲۷۷	-۱/۰۸۶۲	تصادفی
۱۳	-۰/۰۸۸۰	۰/۸۶۴۴	-۰/۱۷۰۶	تصادفی	۳۳	-۰/۳۲۹۹	۰/۱۴۲۵	-۱/۴۶۶۳	تصادفی
۱۴	-۰/۰۵۷۶	۰/۵۳۶۵	۰/۶۱۷۹	تصادفی	۳۴	-۰/۳۶۸۱	۰/۰۹۵۶	-۱/۶۶۶۱	یکنواخت
۱۵	-۰/۱۷۲۸	۰/۱۷۹۹	۱/۳۴۰۹	تصادفی	۳۵	-۰/۲۷۸۷	۰/۲۳۶۹	-۱/۱۸۲۶	تصادفی

خوشه‌هایی که توسط این شاخص کشف می‌شوند در ۷ کلاس دسته‌بندی خواهند شد به این صورت که سه کلاس در دسته Hotspot با سطوح اطمینان ۹۹ درصد، ۹۵ درصد و ۹۰ درصد و سه کلاس در دسته Coldspot با سطوح اطمینان ۹۹ درصد، ۹۵ درصد و ۹۰ درصد قرار می‌گیرند و مناطقی که در هیچکدام از این کلاس‌ها قرار نگیرند در دسته مناطق غیر قابل توجه (بدون اهمیت) قرار می‌گیرند (شکل ۶).

در گام بعد جهت کشف خوشه‌های مکانی بیماری کرونا در این ۶ هفته از تحلیل نقطه داغ استفاده شد که مبنای تشخیص خوشه‌ها در این شاخص براساس مقادیر z-score که همان اختلاف مقادیر از میانگین کل پهنه بوده و در صورتی که این مقدار مثبت باشد نشانگر وجود خوشه با ریسک بالا و پرخطر بوده و در صورتیکه این مقدار منفی باشد نشانگر وجود خوشه با ریسک پایین و کم خطر می‌باشد.

۱ Coldspot



شکل ۶- کشف خوشه‌های مکانی نرخ بروز کرونا با استفاده از شاخص Hotspot Analysis

مکانی، زمانی و مکانی زمانی بیماری‌ها و بررسی معنی‌داری آن‌ها مورد استفاده قرار می‌گیرد. در این نرم‌افزار بایستی سه مجموعه داده شامل تعداد مبتلایان به بیماری در هر شهرستان در طول زمان، جمعیت هر شهرستان، مختصات و موقعیت تقریبی مراکز شهرستان‌ها مشخص شده باشند. این آماره جهت شناسایی و ارزیابی خوشه‌های مبتلایان در طول مکان، زمان و یا مکانی زمانی استفاده می‌شود. این کار با اسکن تدریجی یک پنجره در طول زمان و یا مکان با ذکر تعداد مبتلایان مشاهده شده و موارد مورد انتظار در داخل پنجره در هر موقعیت مکانی انجام می‌شود. این پنجره یک دایره یا بیضی و یا یک استوانه با قاعده دایره‌ای یا بیضی در فضا می‌باشد. اندازه‌های مختلفی برای این پنجره استفاده می‌شود. پنجره با حداکثر احتمال، محتمل‌ترین خوشه است، یعنی خوشه‌ای که کمترین احتمال را داشته باشد از شانس پایینی برای خوشه‌شدن برخوردار خواهد بود. نسبت تعداد موارد بیماری مشاهده‌شده به موارد مورد انتظار نمایانگر معنی‌دار بودن یا نبودن آن خوشه خواهد بود. نتایج این آماره در قالب ۸ کلاس در جدول (۴) و شکل (۷) همراه با خطر نسبی هر کلاس نشان داده شده‌است. از جمله ویژگی‌های ممتاز این روش می‌توان

مطابق نتایج بدست‌آمده برای هفته ۵۱ مربوط به سال ۹۸ شهرستان شبستر به عنوان منطقه با ریسک بالا شناسایی شد. در هفته ۴ سال ۹۹ شهرستان‌های میانه، هشتگرد، اهر، در هفته ۵ شهرستان‌های هریس و کلیبر، در هفته ۱۷ شهرستان‌های هشتگرد، سراب به عنوان مناطق با ریسک بالا و شهرستان کلیبر در این هفته تحت عنوان منطقه با ریسک پایین تشخیص داده شد. در ادامه برای هفته ۱۸ شهرستان‌های بستان‌آباد، سراب به عنوان مناطق با ریسک بالا و شهرستان خدآفرین منطقه با ریسک پایین بودند، در نهایت برای هفته ۱۹ شهرستان‌های بستان‌آباد و هشتگرد مناطق با ریسک بالا تشخیص داده شدند. مطابق نقشه‌های بدست‌آمده خوشه‌ها بیشتر در مناطق شرقی و جنوبی این استان بودند که مناطق با ریسک بالا می‌تواند ناشی از بالا بودن نرخ بروز کرونای هفته ماقبل یا سایر پارامترها باشد و مناطق با ریسک پایین نیز می‌تواند متأثر بر پایین بودن پارامتر نرخ بروز کرونای هفته ماقبل یا سایر متغیرها باشد. در ادامه جهت کشف خوشه‌های مکانی-زمانی بیماری کرونا از آماره اسکن مکانی-زمانی موجود در نرم‌افزار SaTScan که توسط کولدروف توسعه یافته، استفاده شد. این نرم‌افزار جهت کشف خوشه‌های

کلیه هفته‌ها میانگین‌گیری کرده و بیشینه مقدار در بین میانگین‌ها به عنوان پارامترهای تاثیر گذار انتخاب شد که در جدول (۵) قابل مشاهده می‌باشد. ذکر این نکته ضروری است که دلیل پایین بودن همبستگی پیرسون برای برخی پارامترها به دلیل میانگین‌گیری از کلیه هفته‌ها بود.

جدول ۵- نتایج محاسبه VIF و r_{xy} برای پارامترهای منتخب

متغیر وابسته	متغیر مستقل	VIF	r_{xy}
نرخ بروز کرونا	نرخ جمعیت فعال اقتصادی	۲/۱۵۰	۰/۱۶۲۲
	فاصله از اماکن گردشگری	۱/۷۹۶	۰/۱۳۲۱
	فاصله از ترمینال‌ها	۲/۸۹۶	-۰/۱۱۰۹
	نرخ بروز کرونای هفته ماقبل	۲/۰۲۴	۰/۷۳۲۳
	تعداد مبتلایان هر هفته	۱/۹۰۴	۰/۱۲۷
	دمای مینیمم	۲/۲۴۸	-۰/۱۱
	سرعت باد	۲/۰۲۰	-۰/۱۱۳

حدآستانه زمانی برای آموزش هر حالت انجام گرفت به این صورت که شبکه در هر حالت آموزش را تا اتمام حد آستانه زمانی مشخص ادامه داده، سرانجام حالتی که دقت بهتری داشت، انتخاب می‌شود. روند تیون کردن به این شکل انجام گرفت که براساس یک شاخصی که ما در اینجا از شاخص R^2 استفاده کرده‌ایم، تعداد نرون هر لایه را با آزمون و خطا در حالات و تعداد مختلف تست کرده در نهایت براساس شاخص مدنظر حالتی که بهتر از همه بود را انتخاب می‌کنیم هم‌چنین از تابع فعالساز سیگموئید-تانژانت براساس مطالعات گذشته و کارایی بالای آن مورد استفاده قرار گرفت. باتوجه به کمبود داده بیماری کرونا از تابع یادگیری $trainbr^2$ استفاده شد تا بیش‌برازش رخ ندهد. در واقع بیش‌برازش به این موضوع اشاره دارد که مدل ما بسیار خوب آموزش دیده‌است اما به خوبی تعمیم نیافته‌است. ضمناً فرآیند مدلسازی رگرسیون وزندار جغرافیایی در نرم‌افزار ArcGIS 10.7.1 و شبکه عصبی پرسپترون چندلایه در محیط MATLAB 2021 انجام شد.

۳-۴- ارزیابی نتایج و آنالیز حساسیت

جهت ارزیابی نقشه‌های پیش‌بینی نرخ بروز بیماری کرونا در ۸ هفته پایانی (هفته‌های ۲۸ تا ۳۵ مربوط به سال ۱۳۹۹) حاصل از روش‌های MLP و GWR از معیارهایی همچون R^2 ، MAE^3 ، $NRMSE^4$ ، $RMSE^5$ استفاده شد. معیار R^2 آماره است که اطلاعاتی در مورد میزان خوب بودن برازش یک مدل می‌دهد. این معیار بر اساس نسبت تغییرات کل نتایج، میزان قوی بودن رابطه خطی بین دو متغیر را ارزیابی می‌کند که با رابطه (۸) بدست می‌آید:

$$R^2 = \frac{\sum_{i=1}^N (V_i - Y)^2}{\sum_{i=1}^N (Y_i - Y)^2} \quad (8)$$

که V_i مقادیر پیش‌بینی نرخ بروز کرونا، Y_i مقادیر واقعی نرخ بروز کرونا و Y میانگین مقادیر واقعی می‌باشد. مقدار این ضریب بین ۰ و ۱ متغیر بوده که اگر برابر با ۱ باشد نشان‌دهنده تناسب کامل بوده و بنابراین مدل بسیار قابل

۳-۳- مدلسازی

جهت تهیه نقشه‌های پیش‌بینی توزیع مکانی بیماری کرونا به روش رگرسیون وزندار جغرافیایی براساس پارامترهای منتخب جدول (۵) به این صورت عمل شد که براساس کرنل تطبیقی جهت رفع اثرات شهرستان‌های (چندضلعی‌های) غیر همسایه از پارامترهای هفته ماقبل برای پیش‌بینی هفته بعد استفاده شد. هم‌چنین برای مدلسازی به روش شبکه عصبی پرسپترون چندلایه به این شکل عمل شد که مدلسازی بصورت شهرستان به شهرستان بوده و به جهت برقراری شرایط یکسان با روش رگرسیون وزندار جغرافیایی تاخیر زمانی یک هفته‌ای در نظر گرفته شد تا در این روش هم از داده‌های هفته قبل برای پیش‌بینی هفته بعد استفاده شود بعلاوه از ۸۰ درصد داده‌ها (۳۲ هفته) برای آموزش مدل و ۲۰ درصد داده‌ها (۸ هفته) جهت تست و ارزیابی در نظر گرفته شد. در آموزش مدل از یک لایه پنهان همراه با عمل تیون کردن تعداد نرون‌های پهنه لایه پنهان براساس شاخص R^2 و در نظرگرفتن

۱ Goodness of fit

۲ Bayesian regularization backpropagation

۳ Mean absolute error

۴ Normalized Root Mean Square Error

۵ Root Mean Square Error

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (Y_i - \hat{y}_i)^2}{N}} \quad (10)$$

که Y_i مقادیر پیش‌بینی نرخ بروز کرونا، \hat{y}_i مقادیر واقعی نرخ بروز کرونا و N تعداد کل شهرستان‌هایی است که پیش‌بینی می‌شود [۴۰]، [۴۱].

معیار دیگر خطای ریشه میانگین مربع نرمال‌شده (NRMSE) می‌باشد. نرمالسازی RMSE مقایسه بین مجموعه داده‌ها یا مدل‌ها با مقیاس‌ها مختلف را تسهیل می‌کند. این مقدار معمولاً برحسب درصد بیان شده و با استفاده از رابطه (۱۱) قابل محاسبه است:

$$NRMSE = \frac{RMSE}{Y_{obs,max} - Y_{obs,min}} \quad (11)$$

که RMSE همان مقدار بدست‌آمده از رابطه (۹) بوده و $Y_{obs,max}$ بیشینه مقدار مشاهدات واقعی و $Y_{obs,min}$ کمینه مقدار مشاهدات واقعی است [۴۲].

نقشه واقعیت و پیش‌بینی نرخ بروز بیماری کرونا برای ۸ هفته پایانی حاصل از مدل رگرسیون وزندار جغرافیایی در شکل (۸) و نتایج ارزیابی این روش در جدول (۶) نشان داده شده‌اند. همچنین، نقشه‌های حاصل از روش پرسپترون چندلایه در شکل (۹) و نتایج ارزیابی آن در جدول (۸) قابل مشاهده است.

ضرایب حاصل از روش رگرسیون وزندار جغرافیایی برای متغیرهای ورودی در جدول (۷) قابل مشاهده است. ضرایب حاصل برای هر شهرستان متفاوت بوده است. میانگین و انحراف معیار از میانگین‌گیری همه شهرستان‌ها محاسبه شد.

اعتماد برای پیش‌بینی آینده است، در حالی که مقدار صفر نشان می‌دهد که مدل در مدلسازی دقیق داده‌ها شکست خورده است [۳۴]، [۳۵]، [۳۶].

معیار دیگر که در آمار به آن میانگین خطای مطلق (MAE) گفته می‌شود، معیاری از خطاهای بین مشاهدات زوجی است که پدیده یکسانی را بیان می‌کنند. این خطا یک معیار رایج برای خطای پیش‌بینی در تحلیل سری‌های زمانی است و از رابطه (۹) محاسبه می‌شود:

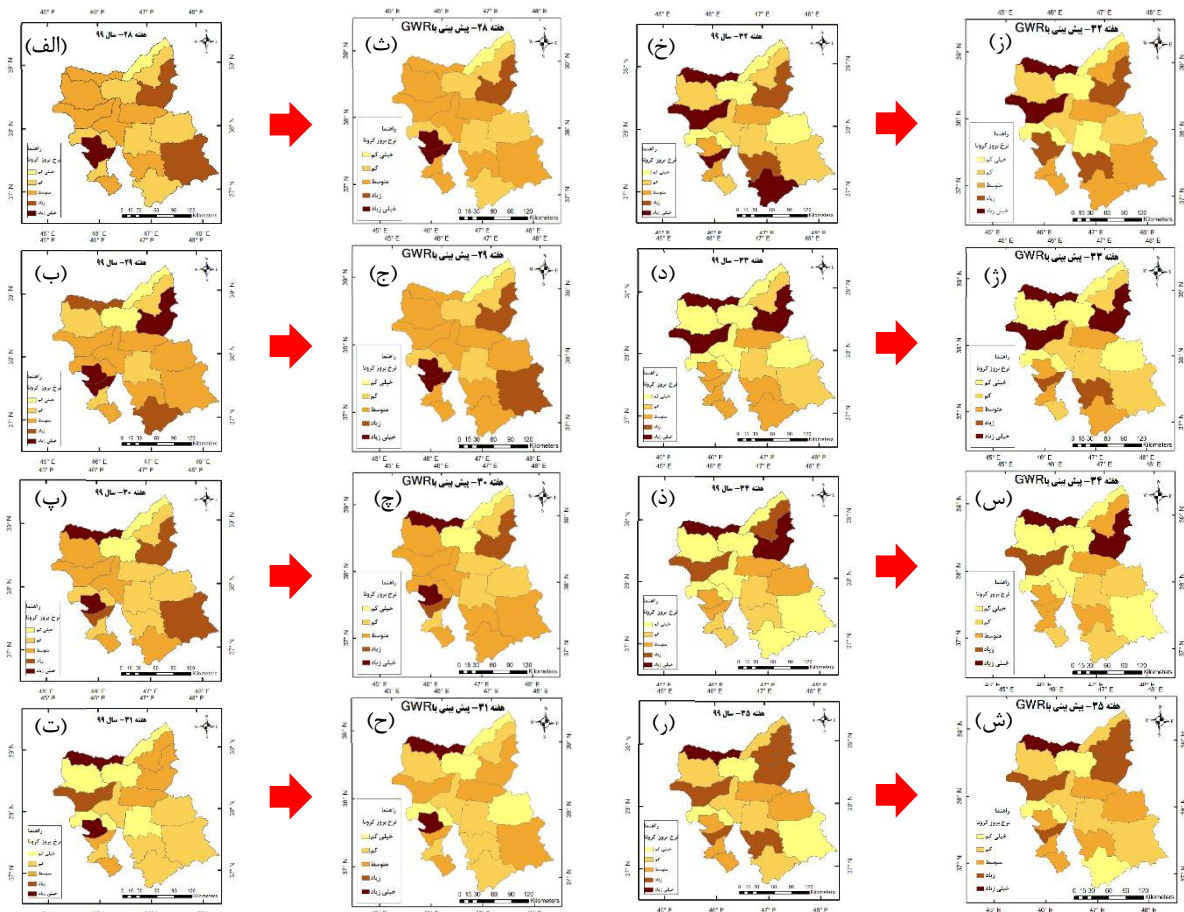
$$MAE = \frac{\sum_{i=1}^N |Y_i - \hat{y}_i|}{N} \quad (9)$$

که Y_i مقادیر پیش‌بینی نرخ بروز کرونا، \hat{y}_i مقادیر واقعی نرخ بروز کرونا و N تعداد کل شهرستان‌هایی است که پیش‌بینی می‌شود [۳۷]، [۳۸].

معیار بعدی خطای جذر میانگین مربعات (RMSE) بوده و معیاری است که اختلاف بین مقادیر مشاهده شده و پیش‌بینی شده توسط یک مدل یا یک برآوردگر را بیان می‌دارد. RMSE یک اندازه‌گیری از دقت بوده که برای مقایسه خطاهای پیش‌بینی مدل‌های مختلف برای یک مجموعه داده خاص و نه بین مجموعه داده‌های مختلف، زیرا وابسته به مقیاس است [۳۹]. RMSE همواره نامنفی بوده و مقدار صفر نشان‌دهنده تناسب کامل با داده‌ها است. بطور کلی هر چه مقدار RMSE کوچکتر باشد بهتر است. این معیار ارزیابی از طریق رابطه (۱۰) قابل محاسبه است:

جدول ۶- نتایج ارزیابی روش رگرسیون وزندار جغرافیایی

هفته	R^2	RMSE	MAE	NRMSE	تعداد موارد درست پیش‌بینی شده	تعداد موارد اشتباه پیش‌بینی شده	درصد تناظر صحیح
۲۸	۰/۹۲۴۱	۰/۱۲۴۱	۰/۱۰۲۵	۰/۰۶۲	۱۷	۲	۸۹/۴۷
۲۹	۰/۸۷۸۶	۰/۱۷۹۶	۰/۱۳۳۳	۰/۰۸۹۸	۱۱	۸	۵۷/۸۹
۳۰	۰/۹۲۴۳	۰/۱۴۰۶	۰/۱۱۸۷	۰/۰۷۰۳	۱۶	۳	۸۴/۲۱
۳۱	۰/۹۱۱۲	۰/۱۴۱۴	۰/۱۰۹۹	۰/۰۷۰۷	۱۳	۶	۶۸/۴۲
۳۲	۰/۷۷۳۱	۰/۲۷۳۳	۰/۲۲۱۴	۰/۱۳۶۶	۱۱	۸	۵۷/۸۹
۳۳	۰/۹۳۲۸	۰/۱۴۶۶	۰/۱۲۱۷	۰/۰۷۳۳	۱۵	۴	۷۸/۹۴
۳۴	۰/۸۹۴۱	۰/۱۹۴۵	۰/۱۴۷۴	۰/۰۹۷۲	۱۴	۵	۷۳/۶۸
۳۵	۰/۹۵۰۱	۰/۱۱۶۲	۰/۰۹۰۸	۰/۰۵۸۱	۱۶	۳	۸۴/۲۱



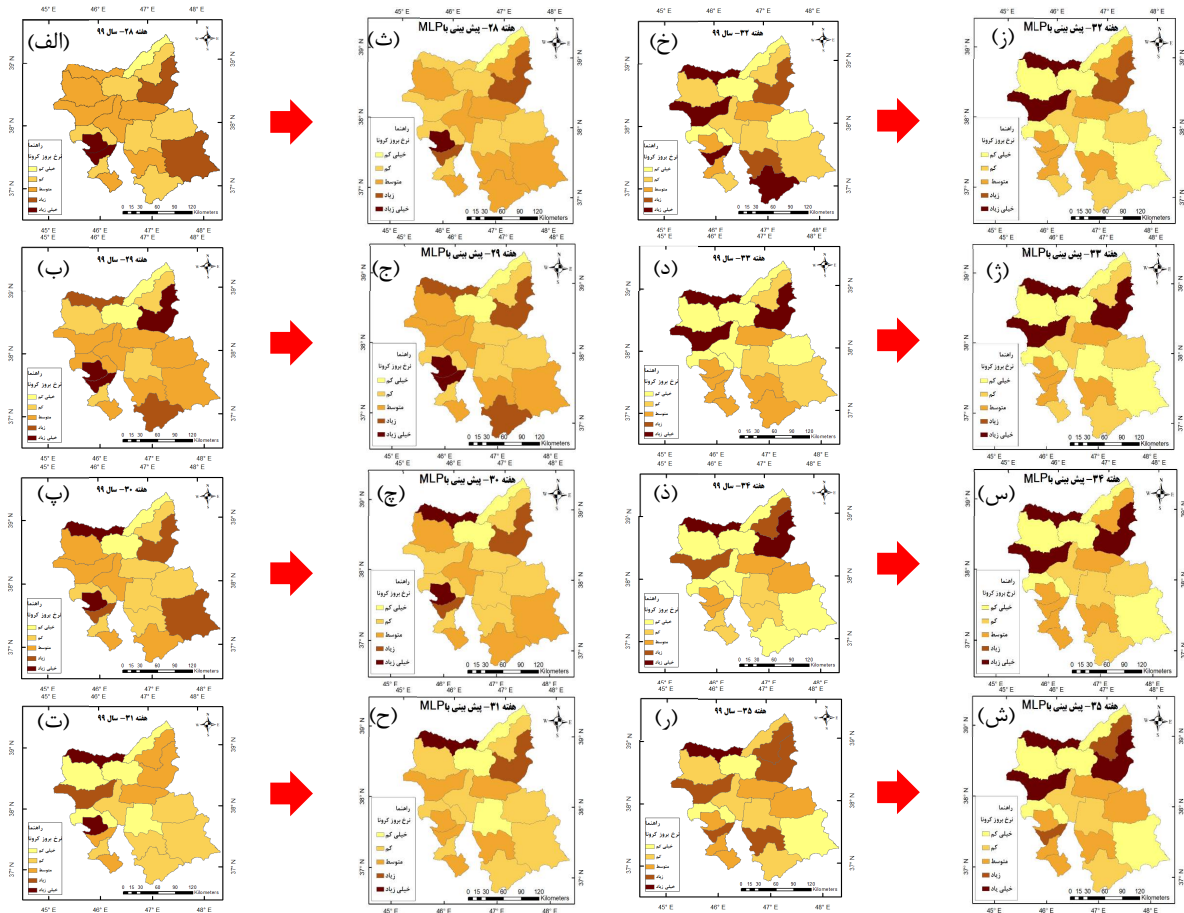
شکل ۸- (الف،ب،ت) نقشه‌های واقعیت نرخ بروز بیماری کرونای هفته‌های ۲۸ تا ۳۱. (ث،ج،ح) نقشه‌های پیش‌بینی هفته‌های ۲۸ تا ۳۱ با استفاده از GWR. (خ،د،ذ،ر) نقشه‌های واقعیت هفته‌های ۲۲ تا ۲۵. (ز،ژ،س،ش) نقشه‌های پیش‌بینی هفته‌های ۲۲ تا ۲۵ با استفاده از GWR.

جدول ۷- ضرایب بدست‌آمده از روش رگرسیون وزندار جغرافیایی

پارامتر ورودی	حداقل	حداکثر	میانگین	انحراف معیار
نرخ جمعیت فعال اقتصادی	۰/۰۰۹۶	۰/۵۵۴۴	۰/۲۰۹۴	۰/۲۶۸۴
فاصله از سایت‌های گردشگری	۰/۰۰۷۱	۰/۶۷۱۹	۰/۲۰۵۲	۰/۲۰۱۱
فاصله از ترمینال‌ها	۰/۰۰۰۲	۰/۵۴۲۹	۰/۱۵۷۴	۰/۳۲۴۷
تعداد میتلایان	۰/۰۰۰۴	۰/۳۲۲۸	۰/۱۲۰۱	۰/۲۵۲۶
دمای کمینه	۰/۰۰۲	۰/۴۹۹۶	۰/۱۶۳۲	۰/۲۰۱۷
سرعت باد	۰/۰۰۰۸	۰/۴۵۴۲	۰/۱۰۸۷	۰/۲۶۴۵
نرخ بروز کرونای هفته ما قبل	۰/۴۰۳۸	۱/۷۳۷۹	۰/۸۲۳	۰/۲۴۰۰

جدول ۸- نتایج ارزیابی روش شبکه عصبی پرسپترون چندلایه

هفته	R^2	RMSE	MAE	NRMSE	تعداد پیش‌بینی درست	تعداد پیش‌بینی اشتباه	درصد تناظر صحیح
۲۸	۰/۷۷۶۲	۱۱۵/۲۵۵۱	۸۹/۶۰۷۵	۰/۱۲۱۴	۱۳	۶	۶۸/۴۲
۲۹	۰/۸۷۶۷	۱۰۶/۲۰۸	۸۹/۹۸۶۱	۰/۱۲۹۱	۱۴	۵	۷۳/۶۸
۳۰	۰/۸۷۹۲	۱۴۵/۰۷۱۵	۱۱۱/۴۸۹۳	۰/۱۱۳۲	۱۷	۲	۸۹/۴۷
۳۱	۰/۸۹۳۳	۱۱۱/۸۹۹۷	۸۳/۶۷	۰/۰۹۱۰	۱۲	۷	۶۳/۱۵
۳۲	۰/۶۵۲۴	۲۳۹/۸۱۲۵	۱۹۸/۰۲۰۴	۰/۲۰۹۱	۱۱	۸	۵۷/۸۹
۳۳	۰/۹۳۱۲	۱۵۷/۳۸۴۵	۱۲۱/۱۹۰۶	۰/۰۸۷۳	۱۳	۶	۶۸/۴۲
۳۴	۰/۸۰۴	۳۳۸/۱۶۶۸	۲۲۰/۸۲۲۶	۰/۱۶۴۵	۱۱	۸	۵۷/۸۹
۳۵	۰/۷۶۸۱	۲۴۴/۳۲	۱۵۷/۰۸۵۱	۰/۱۵۶۹	۱۴	۵	۷۳/۶۸



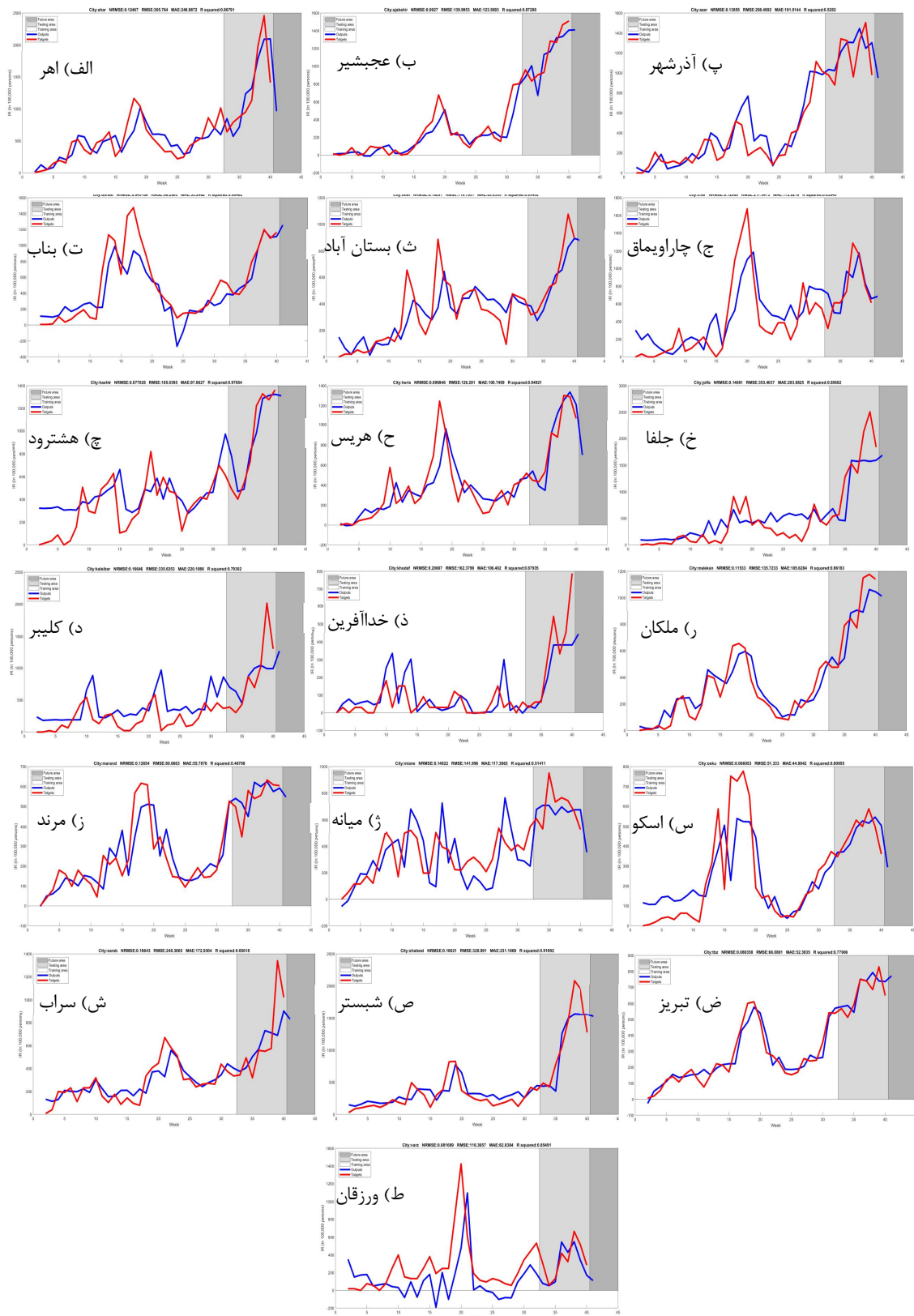
شکل ۹- (الف، ب، پ، ت) نقشه‌های واقعیت نرخ بروز بیماری کرونای هفته‌های ۲۸ تا ۳۵. (ث، ج، ح، ش) نقشه‌های پیش‌بینی هفته‌های ۲۸ تا ۳۵ با استفاده از MLP. (خ، ذ، ر) نقشه‌های واقعیت هفته‌های ۳۲ تا ۳۵. (ز، ژ، س، ش) نقشه‌های پیش‌بینی هفته‌های ۳۲ تا ۳۵ با استفاده از MLP.

هفته‌های ۲۸ تا ۳۵ می‌باشند، در نمودارها همان هفته‌های ۳۳ تا ۴۰ خواهند بود، البته ذکر این نکته ضروری است که در جداول این شماره‌ها تغییری نکرده و همان شماره‌های قبلی هستند. محور عمودی نیز بیانگر نرخ بروز کرونا در مقیاس ۱۰۰۰۰۰ نفر می‌باشد. در این نمودارها هرچه نمودار آبی (مقادیر پیش‌بینی) به نمودار قرمز رنگ (مقادیر واقعی) نزدیکتر باشد نشان‌دهنده نزدیک بودن مقدار پیش‌بینی به مقدار واقعی در آن هفته‌ها است.

باتوجه به اینکه در روش شبکه عصبی پرسپترون چندلایه برخلاف روش رگرسیون وزن‌دار جغرافیایی از نحوه وزن‌دهی و ضرایب مورد استفاده در آن آگاهی نداریم، به عبارتی به شکل جعبه سیاه عمل می‌کند، بنابراین بررسی اهمیت نسبی پارامترهای ورودی امکان‌پذیر نیست. با این حال، از تحلیل حساسیت می‌توان برای بررسی سهم پارامترهای ورودی در مدل‌سازی و پیش‌بینی استفاده کرد.

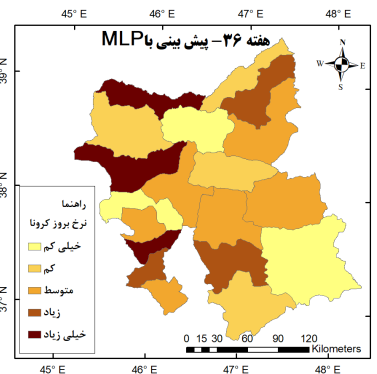
در مدل‌سازی به روش MLP نمودارهای مقایسه مقادیر واقعی (رنگ قرمز) و پیش‌بینی (رنگ آبی) نرخ بروز کرونا برای هر شهرستان به صورت جداگانه در شکل (۱۰) نشان داده شده‌است. با توجه به شکل (۱۰) هر نمودار حاوی سه ناحیه آموزش، تست و ارزیابی و آینده می‌باشد. ناحیه سفید (آموزش) حاوی ۳۲ هفته، ناحیه نقره‌ای (تست و ارزیابی) حاوی ۸ هفته و ناحیه خاکستری (آینده) می‌باشد که باتوجه به تاخیر زمانی یک هفته‌ای در نظر گرفته شده در مدل‌سازی MLP، شبکه قادر خواهد بود تا تنها یک هفته بعد از آخرین هفته از داده‌های موجود را پیش‌بینی نماید که نقشه پیش‌بینی این هفته در شکل (۱۱) نشان داده شده‌است.

محور افقی در شکل (۱۰) نشانگر شماره هفته بوده که در این جا پنج هفته ابتدایی که مربوط به اواخر سال ۹۸ بود، از یک شروع شده و به تبع به انتهای داده‌های ما نیز ۵ هفته اضافه خواهد شد عبارتی بطور مثال، داده‌های تست ما که



شکل ۱۰- نمودارهای مقایسه مقادیر واقعی و پیش‌بینی نرخ بروز کرونا برای هر شهرستان

پارامتر از مدل خارج شده و تأثیر آن پارامتر بر عملکرد مدل بر اساس معیارهای ارزیابی تعیین می‌شود. متغیری که بیشترین کاهش را در کارایی مدل داشته باشد به عنوان تاثیرگذارترین متغیر انتخاب می‌شود. میانگین مقادیر بدست‌آمده از آنالیز حساسیت متغیرها در جدول (۹) نشان داده شده‌است. شایان ذکر است که آنالیز حساسیت برای روش رگرسیون وزندار جغرافیایی نیز پیاده سازی شده‌است. نتایج بدست‌آمده از جدول (۷) با نتایج حاصل از جدول (۹) در رابطه با تعیین مهم‌ترین متغیر تاثیرگذار در مدلسازی به روش رگرسیون وزندار جغرافیایی همخوانی دارد.



شکل ۱۱- نقشه پیش‌بینی هفته ۳۶ با استفاده از MLP

برای انجام تجزیه و تحلیل حساسیت، در هر اجرا یک

جدول ۹- نتایج حاصل از آنالیز حساسیت پارامترهای موثر در مدلسازی

پارامتر ورودی	مدل رگرسیون وزندار جغرافیایی (GWR)				مدل شبکه عصبی پرسپترون چندلایه (MLP)			
	NRMSE	MAE	RMSE	R ²	NRMSE	MAE	RMSE	R ²
نرخ جمعیت فعال اقتصادی	۰/۰۹۲	۰/۱۴۵۹	۰/۱۸۴۱	۰/۸۷۵۳	۰/۱۵۶۴	۱۶۳/۴۰۵۹	۲۱۷/۷۹۹۱	۰/۷۴۹۷
فاصله از سایت‌های گردشگری	۰/۰۹۱۷	۰/۱۴۷۰	۰/۱۸۳۵	۰/۸۷۱۱	۰/۱۶۵۳	۱۵۷/۰۹۴۴	۲۱۵/۷۹۰۸	۰/۶۹۳۹
فاصله از ترمینال‌ها	۰/۰۸۷۶	۰/۱۳۷۴	۰/۱۷۵۲	۰/۸۸۵۶	۰/۱۵۷۲	۱۶۲/۸۳۹۶	۲۱۴/۴۸۳۷	۰/۷۶۴۷
تعداد مبتلایان	۰/۰۷۸۲	۰/۱۴۲۱	۰/۱۵۶۶	۰/۸۸۸۵	۰/۱۴۸۲	۱۴۸/۱۴۶۲	۲۰۰/۸۴۰۷	۰/۷۸۴۱
دمای کمینه	۰/۰۸۸۳	۰/۱۴۰۲	۰/۱۷۶۷	۰/۸۸۴۰	۰/۱۴۴۷	۱۵۳/۳۴۹۴	۲۰۰/۴۳۱۹	۰/۷۸۸۵
سرعت باد	۰/۰۸۱۷	۰/۱۴۱۵	۰/۱۶۳۵	۰/۸۸۱۰	۰/۱۸۴۳	۱۷۹/۳۱۸۵	۲۵۶/۶۲۲۹	۰/۶۳۶۹
نرخ بروز کرونای هفته قبل	۰/۱۶۴۸	۰/۲۳۹۸	۰/۳۲۵۸	۰/۶۱۲۵	۰/۱۶۷۵	۱۷۴/۰۶۲۱	۲۳۰/۶۰۵	۰/۷۱۳۱

همچون حالت کلی در همه شهرستان‌ها متغیرهای نرخ بروز هفته ماقبل و تعداد مبتلایان هر شهرستان به عنوان پارامترهای موثر در بروز این بیماری شناخته شدند که این نشان‌دهنده میزان بالا بودن انتقال بیماری از انسان به انسان می‌باشد. هم‌چنین، در چند شهرستان علاوه بر این دو، پارامترهای دیگری همچون تعداد فوتی‌های هر شهرستان، سرعت باد، تابش خورشیدی و رطوبت نسبی به عنوان پارامترهای تاثیرگذار در بروز این بیماری برای آن‌ها شناسایی شدند. بطور متوسط مقادیر بدست‌آمده برای معیارهای ارزیابی در مدل MLP برای $R^2=0/8711$ و $NRMSE=0/8319$ بودند که در مقایسه با حالت کلی مدل MLP، بر مبنای معیار R^2 دقت کار افزایش یافته است که این نشان می‌دهد مدلسازی محلی مناسب‌تر است.

در ادامه این تحقیق علاوه بر تعیین پارامترهای موثر به صورت کلی، پارامترهای موثر را به تفکیک شهرستان شناسایی و با مدل MLP روند مدلسازی را تکرار کردیم. روش کار به این صورت بود که مطابق داده‌های در نظر گرفته شده در جدول (۱) برای هر شهرستان و نرخ بروز کرونای آن‌ها میزان همبستگی آن‌ها را بدست آورده و پس از مشخص شدن پارامترهای موثر هر شهرستان فرآیند مدلسازی را با مدل MLP برای هر شهرستان به صورت جداگانه تکرار می‌کنیم که نتایج مربوطه به همراه مقادیر بدست‌آمده برای شاخص‌های ارزیابی R^2 و $NRMSE$ در جدول (۱۰) نشان داده شده‌است. در اینجا نیز نسبت‌های مختلف برای تعیین اندازه مناسب برای داده‌های آموزش و تست و ارزیابی انجام شد و مشخص گردید که میزان ۷۰ درصد برای آموزش و ۳۰ درصد برای تست و ارزیابی اندازه‌های مناسبی هستند. مطابق نتایج بدست‌آمده

جدول ۱۰- مدل‌سازی براساس پارامترهای موثر منتخب هر شهرستان

شهرستان	پارامتر موثر (مقدار همبستگی پیرسون (I_{xy}))	NRMSE	R ²	RMSE	MAE
آذرشهر	نرخ بروز کرونای هفته ما قبل و تعداد مبتلایان هر هفته (۰/۰۹) - تعداد مرگ و میر هر هفته (۰/۵۴۲)	۰/۱۴۶۱	۰/۷۷۷۹	۲۱۹/۹۶	۱۹۵/۹۱
اسکو	نرخ بروز کرونای هفته ما قبل و تعداد مبتلایان هر هفته (۰/۷۹)	۰/۰۷۶۶	۰/۸۹۶۵	۵۹/۵۷	۴۷/۱۴
اهر	نرخ بروز کرونای هفته ما قبل و تعداد مبتلایان هر هفته (۰/۸۲۵) - تعداد مرگ و میر هر هفته (۰/۶۸۴)	۰/۱۸۲۳	۰/۹۳۷۸	۴۴۷/۳۳	۳۱۶/۸۴
بستان آباد	نرخ بروز کرونای هفته ما قبل و تعداد مبتلایان هر هفته (۰/۷۵۶)	۰/۱۵۳۹	۰/۸۳۵۶	۱۶۵/۶۵	۱۳۸/۶
بناب	نرخ بروز کرونای هفته ما قبل و تعداد مبتلایان هر هفته (۰/۸۶۶)	۰/۰۷۹۶	۰/۹۳۸۸	۱۱۶/۹۷	۹۴/۰۸
تبریز	نرخ بروز کرونای هفته ما قبل و تعداد مبتلایان هر هفته (۰/۹۲۹) - تعداد مرگ و میر هر هفته (۰/۷۴۶)	۰/۱۴۸۵	۰/۸۹۳۳	۱۲۲/۱۷	۹۹/۶۵
جلفا	نرخ بروز کرونای هفته ما قبل و تعداد مبتلایان هر هفته (۰/۸۸۵) - تعداد مرگ و میر هر هفته (۰/۶۰۷)	۰/۲۳۰۳	۰/۸۶۱۶	۵۷۸/۱۹	۴۵۰/۸۶
خداآفرین	نرخ بروز کرونای هفته ما قبل و تعداد مبتلایان هر هفته (۰/۷۸۲) - سرعت باد (۰/۵۴۵) - تابش خورشیدی (۰/۵۰۳)	۰/۲۶۸۷	۰/۸۸۴۶	۲۱۱/۷۹	۱۳۸/۴۷
سراب	نرخ بروز کرونای هفته ما قبل و تعداد مبتلایان هر هفته (۰/۷۷)	۰/۱۹۶۳	۰/۶۸۷۶	۲۶۱/۵۴	۱۵۶/۲۸
شبستر	نرخ بروز کرونای هفته ما قبل و تعداد مبتلایان هر هفته (۰/۸۹۴) - تعداد مرگ و میر هر هفته (۰/۶۳۷)	۰/۲۷۹۹	۰/۹۳۴۵	۵۷۴/۶۴	۳۹۴/۵۸
عجبشیر	نرخ بروز کرونای هفته ما قبل و تعداد مبتلایان هر هفته (۰/۹۵۶) - تعداد مرگ و میر هر هفته (۰/۷۱۳)	۰/۱۹۷۵	۰/۸۷۱	۲۹۸/۳۴	۲۴۰/۹
کلیبر	نرخ بروز کرونای هفته ما قبل و تعداد مبتلایان هر هفته (۰/۷۶۵)	۰/۲۴۳	۰/۸۰۴۸	۴۹۰/۰۴	۱۰۷/۲۴
مرند	نرخ بروز کرونای هفته ما قبل و تعداد مبتلایان هر هفته (۰/۸۳۹)	۰/۱۹۴۲	۰/۸۴۲	۱۲۲/۹۳	۱۰۳/۶۱
ملکان	نرخ بروز کرونای هفته ما قبل و تعداد مبتلایان هر هفته (۰/۹۲۴) - تعداد مرگ و میر هر هفته (۰/۵۹۷)	۰/۱۹۷	۰/۹۳۰۵	۲۳۱/۸۹	۱۷۸/۳۳
میانه	نرخ بروز کرونای هفته ما قبل و تعداد مبتلایان هر هفته (۰/۷۶۷)	۰/۱۵۸۹	۰/۶۱۳۲	۱۵۱/۳	۱۰۷/۲۴
هریس	نرخ بروز کرونای هفته ما قبل و تعداد مبتلایان هر هفته (۰/۸۳۴) - تعداد مرگ و میر هر هفته (۰/۶۰۹)	۰/۰۹۳۷	۰/۹۴۶۸	۱۲۲/۱۴	۹۸/۸۱
هشترود	نرخ بروز کرونای هفته ما قبل و تعداد مبتلایان هر هفته (۰/۸۵۳)	۰/۱۴۲۲	۰/۹۲۱۷	۱۹۳/۹۹	۱۶۲/۱۶
ورزقان	نرخ بروز کرونای هفته ما قبل و تعداد مبتلایان هر هفته (۰/۶۱۸)	۰/۱۱۳۹	۰/۵۸۴۶	۱۶۲/۲۹	۱۲۶/۲
چاراویماق	نرخ بروز کرونای هفته ما قبل و تعداد مبتلایان هر هفته (۰/۷۶۷) - رطوبت نسبی (۰/۵۲۱)	۰/۱۴۷۲	۰/۶۴۳۱	۲۴۶/۴۹	۱۹۷/۸۹

۳-۵- بحث

سلامت و بهداشت، اقتصادی و تسهیلات شهری مختلفی در نظر گرفته شد. مطابق نتایج حاصل از ضریب تورم واریانس وابستگی قابل ملاحظه‌ای بین متغیرها با نرخ بروز کرونا مشاهده نشد و فرض مستقل بودن پارامترها فرض درستی بود. براساس نتایج حاصل از ضریب همبستگی پیرسون با صرف نظر از علامت مثبت و منفی بترتیب، پارامترهای نرخ

در این مطالعه از روش‌های رگرسیون وزندار جغرافیایی و شبکه عصبی پرسپترون چندلایه جهت دستیابی به نقشه‌های پیش‌بینی توزیع مکانی-زمانی نرخ بروز بیماری کرونا در استان آذربایجان شرقی استفاده شده‌است. در این مدل‌سازی پارامترهای مختلفی همچون پارامترهای محیطی، توپوگرافی،

ارجحیت داشت و در یک هفته هم هر دو تعداد شهرستان یکسانی را پیش‌بینی کرده بودند؛ بنابراین، مدل GWR نسبت به مدل MLP تعداد هفته‌ی بیشتری را درست پیش‌بینی کرده بود. همچنین، براساس معیار ارزیابی NRMSE و R^2 بطور میانگین برای مدل GWR مقدار $R^2 = 0/8985$ و $NRMSE = 0/0822$ و برای مدل MLP نیز $R^2 = 0/8226$ و $NRMSE = 0/134$ بدست آمد که بیانگر این است که روش GWR مناسب‌تر است. در این تحقیق، علاوه بر تعیین پارامترهای موثر بروز کرونا به صورت کلی، به تفکیک شهرستان نیز انجام و پارامترهای موثر هر شهرستان استخراج گردید که برای تمام شهرستان‌ها متغیرهای نرخ بروز هفته ماقبل و تعداد مبتلایان به عنوان پارامترهای تاثیر گذار شناخته شد. سپس، براساس پارامترهای موثر هر شهرستان فرآیند مدلسازی به روش MLP تکرار گردید و یک مدل محلی ارائه شد که با مقایسه نتایج ارزیابی مشخص شد که دقت مدل محلی بهتر از حالت کلی روش MLP است. باتوجه به نمودارهای مقایسه مقادیر واقعی و پیش‌بینی نرخ بروز کرونا روند بروز این بیماری در شهرستان‌های اهر، تبریز، سراب، شبستر، عجبشیر، مرند، ملکان و هریس شبیه به هم بوده و دو پیک تشخیص داده شد. در این نمودارها نیز واضح بوده که پیک دوم نسبت به پیک اول شدیدتر بود. همچنین، روند بروز بیماری در شهرستان‌های اسکو، بستان آباد و بناب شبیه هم بوده که در مقطعی دوبار در فاصله کمی به اوج رسید سپس کاهش یافته و بعد دوباره افزایش یافت؛ شهرستان‌های میانه و هشترود نیز شبیه به هم بوده بطوریکه دائماً در بازه زمانی مورد نظر نرخ بروز بیماری در حال افزایش و کاهش بوده و عبارتی پرنوسان بودند؛ شهرستان‌های جلفا، خداآفرین و کلیبر هم روند بروز بیماری شبیه به هم داشتند طوری که تقریباً بعد از هفته ۲۸ نرخ بروز بیماری کرونا در این شهرستان‌ها با شیب تندی افزایش یافت. مطابق نتایج حاصل از آنالیز حساسیت در مدل GWR پارامتر نرخ بروز کرونا هفته ماقبل و در مدل MLP پارامتر سرعت باد با توجه به اینکه کارایی مدل‌ها را بیشتر از سایر پارامترها کاهش می‌دهند و بیشترین تاثیر را بروی مدل‌ها دارند؛ لذا به عنوان مهم‌ترین متغیرهای موثر انتخاب شدند که براساس پارامتر نرخ بروز کرونا هفته ماقبل می‌توان به این نکته پی برد که انتقال بیماری کرونا از انسان به انسان بسیار اتفاق می‌افتد و بایستی خطر ابتلا به این بیماری توسط افراد جدی گرفته شده و پروتکل‌های بهداشتی بیش از پیش رعایت گردد.

بروز بیماری کرونای هفته ماقبل، نرخ جمعیت فعال اقتصادی، فاصله از سایت‌های گردشگری، سرعت باد، دمای کمینه و تعداد مبتلایان به عنوان موثرترین عوامل بروز بیماری کرونا بودند بطوریکه براساس علامت ضرایب بدست‌آمده نرخ بروز بیماری کرونا با عوامل نرخ بروز کرونا هفته ماقبل، نرخ جمعیت فعال اقتصادی، فاصله از سایت‌های گردشگری و تعداد مبتلایان رابطه مستقیم و با عوامل دمای کمینه، سرعت باد و فاصله از ترمینال‌ها رابطه غیرمستقیم وجود داشت. تحلیل‌های زمین آمار نرخ بروز بیماری کرونا بصورت هفتگی نشان داد که تنها در شش هفته خود همبستگی مکانی معنی‌داری وجود دارد و نشان داد که تنها در این شش هفته الگوی توزیع نرخ بروز بیماری بصورت خوشه‌ای بوده و درمابقی هفته‌ها الگوی توزیع نرخ بروز بیماری خوشه‌ای نبود که برای یافتن خوشه‌های مکانی نرخ بروز کرونا در این شش هفته از تحلیل نقطه داغ استفاده کردیم که نتایج آن در شکل (۶) نشان داده شده است که مطابق نقشه‌های بدست‌آمده خوشه‌ها بیشتر در مناطق شرق و جنوبی این استان بودند که مناطق با ریسک بالا می‌تواند ناشی از بالا بودن نرخ بروز کرونای هفته قبل یا سایر پارامترها باشد و مناطق با ریسک پایین نیز می‌تواند متأثر بر پایین بودن پارامتر نرخ بروز کرونای هفته قبل یا سایر متغیرها باشد. جالب توجه است که در این ۶ هفته برخی مناطق با اینکه در همسایگی مناطق کم‌خطر قرار دارند به عنوان مناطق پرخطر شناسایی شدند جهت علت‌یابی این موضوع نیازمند تحقیقات بیشتر و درنظر گرفتن سایر عوامل هستند. سپس، جهت کشف خوشه‌های مکانی-زمانی بیماری از نرم افزار SaTScan استفاده شد. درگام بعد جهت مدلسازی و پیش‌بینی بیماری کرونا با استفاده از GWR و MLP بدین صورت عمل شد که از داده‌های هفته قبل هر هفته، مطابق پارامترهای منتخب همبستگی پیرسون (جدول ۵) استفاده شد. مطابق مقادیر پیش‌بینی شده برای نرخ بروز بیماری کرونا در هر شهرستان توسط هر دو مدل MLP و GWR و کلاسه‌بندی نقشه‌ها به منظور شناسایی مناطق پرخطر برای ۸ هفته (۲۰ درصد داده‌ها) مربوط به تست و ارزیابی انجام شد که اگر تعداد شهرستان‌هایی که کلاس واقعی و پیش‌بینی آن‌ها یکی بود را مبنای مقایسه قرار دهیم، نشان می‌دهد که در پنج هفته مدل GWR نسبت به MLP ارجحیت بالاتری داشته و تعداد شهرستان بیشتری را نسبت به MLP درست پیش‌بینی کرده بود هم چنین در دو هفته دیگر مدل MLP نسبت به GWR

همچنین، می‌توان استنباط نمود که با افزایش سرعت باد ذرات بسیار ریز کروناویروس معلق در هوا به سایر مناطق انتشار یافته که منجر به افزایش احتمال ابتلای افراد حاضر در آن مناطق خواهد شد.

۴- نتیجه‌گیری و پیشنهادات

براساس آمار کل مبتلایان هفتگی بیماری کرونا در استان آذربایجان شرقی در طول ۴۰ هفته در بازه زمانی ۳ اسفند ۱۳۹۸ تا ۳۰ آبان ۱۳۹۹ به نظر می‌رسد که روند شیوع بیماری کرونا طی هفته‌های مختلف تغییرات محسوسی داشته و به این صورت بود که دو پیک را پشت‌سر گذاشته که پیک اول در هفته‌های ۱۳ و ۱۴ و پیک دوم در هفته ۳۴ به اوج خود رسیدند. مطابق نتایج بدست‌آمده از شاخص عمومی موران در طول ۴۰ هفته تنها در شش هفته (هفته‌های ۵۱ مربوط به سال ۱۳۹۸، هفته‌های ۵، ۴، ۱۷، ۱۸، ۱۹ مربوط به سال ۱۳۹۹) بدلیل نزدیک‌بودن مقدار شاخص موران به عدد ۱ الگوی توزیع بیماری خوشه‌ای بوده و فرض صفر در این هفته‌ها رد شد اما در بقیه هفته‌ها الگوی توزیع بیماری تصادفی و یا یکنواخت بود سپس به کشف خوشه‌ها در این ۶ هفته پرداختیم که نشان داد بروز کرونا بیشتر در مناطق شرق و جنوبی این استان متمرکز بود. براساس نتایج حاصل از ضریب تورم واریانس مستقل بودن پارامترهای در نظر گرفته شده فرض درستی بود و می‌توان آن‌ها را در مدل‌سازی دخالت داد. مطابق نتایج حاصل از همبستگی پیرسون پارامترهای نرخ جمعیت فعال اقتصادی، فاصله از اماکن گردشگری، فاصله از ترمینال‌ها، نرخ بروز کرونای هفته ماقبل، تعداد مبتلایان هر هفته، دمای کمینه و سرعت باد به عنوان پارامترهای موثر در بروز کرونا مشخص شدند. در ادامه به منظور مدل‌سازی و پیش‌بینی مناطق پرخطر این بیماری براساس پارامترهای منتخب از مدل‌های رگرسیون وزندار جغرافیایی و شبکه عصبی پرسپترون به این صورت استفاده شد که از داده‌های هفته ماقبل هر هفته در مدل‌سازی و پیش‌بینی آن استفاده شد. مطابق نتایج حاصل از دو مدل براساس تعداد مواردی که هر مدل درست پیش‌بینی کرده بودند، در برخی هفته‌ها مدل GWR به عنوان مدل برتر و در برخی هفته‌ها هم مدل MLP مدل برتر بود اما براساس معیارهای ارزیابی NRMSE و R^2 مدل GWR در مقایسه با مدل MLP مدل برتری شناخته شد. مطابق نتیجه حاصل از پیش‌بینی آینده (هفته ۳۶) شهرستان‌های جلفا، شبستر و

عجشیر به عنوان مناطق پرخطر شناسایی شدند. براساس نتایج بدست‌آمده از تحلیل حساسیت متغیر نرخ بروز هفته ماقبل به عنوان موثرترین پارامتر در مدل‌سازی بیماری کرونا به روش GWR و متغیر سرعت باد به عنوان موثرترین پارامتر در مدل‌سازی بیماری کرونا به روش MLP شناخته شدند. همچنین، پارامترهای موثر به تفکیک هر شهرستان استخراج گردید و یک مدل محلی ارائه شد که در مقایسه با حالت کلی، مدل محلی دقت بهتری نسبت به مدل کلی در روش MLP داشت. به جهت انتقال این بیماری از انسان به انسان بایستی پروتکل‌های بهداشتی بیشتر رعایت گردد. همچنین، مسئولان تصمیم‌گیر به منظور رعایت هر چه بیشتر این بیماری در جامعه به خصوص در مناطقی که به عنوان مناطق پرخطر شناخته شدند، تدابیری اندیشیده و محدودیت‌های لازم را اجرا نمایند تا خطر ابتلا شدن به این بیماری منحوس کاهش یابد. باتوجه به تحقیقات انجام‌شده و مقالاتی که مطالعه شد، مدل‌سازی پیش‌بینی در مسیر تبدیل شدن به یکی از حرفه‌های تجزیه و تحلیل مکانی-زمانی در گسترش بیماری کرونا بوده و دانستن پویایی مکانی و زمانی بیماری کرونا برای کاهش ابتلا و مدیریت آن بسیار مهم است که لازمه آن شناخت مهم-ترین فاکتورهای موثر برای هر بیماری همه‌گیر از جمله کرونا و توسعه سریع اطلاعات قوی برای ارائه آگاهی موقعیتی و بهبود پاسخ به بیماری همه‌ساله بسیار حیاتی است. با استفاده از چنین نقشه‌هایی می‌توان مناطق مستعد و در معرض خطر بالای بیماری کرونا را شناخته و مکان‌های الویت‌دار در افزایش مراقبت‌های بهداشتی برای جمعیت آسیب‌پذیر را شناسایی کرده و به تصمیم‌گیران در راستای ایجاد محدودیت‌ها و افزایش پروتکل‌های بهداشتی کمک کرد. در مطالعه ما چندین محدودیت وجود دارد که از جمله آن‌ها می‌توان به عدم دسترسی به موقعیت مکانی بیماران (داده‌های نقطه‌ای) اشاره کرد که در صورت دسترسی و تهیه چنین داده‌هایی می‌توان مطالعات بیشتر و دقیقتر در راستای کمک به تصمیم‌گیران ارائه کرد. از دیگر محدودیت‌ها می‌توان به محدودیت زمانی و داده‌ای اشاره کرد که در این مطالعه بصورت هفتگی بود و می‌توان در مطالعات آینده بصورت روزانه و در سطح بخش و یا روستا در محدودیت زمانی بیشتر با در نظر گرفتن معیارهای دیگر انجام داد. به منظور تحقیق و مطالعه بیشتر در این زمینه توصیه می‌شود از روش‌هایی همچون مدل‌سازی موجک، دیگر مدل‌های شبکه‌های عصبی همچون LSTM، SVM، NARX، و RBF و روش استنتاج بیسین (BMP) برای درک شیوع و

بیماری و پارامترهای بیشتر و دیگر عوامل تاثیرگذار انتخاب نشده در مقایسه با سایر کشورها و مطالعات، مورد بررسی قرار گیرد.

پیش‌بینی موارد کووید ۱۹ استفاده گردد که از مزیت‌های آن می‌توان به ترکیب چندین پارامتر برای پیش‌بینی آماری قوی اشاره کرد. همچنین، می‌توان با در نظر گرفتن دوره نهفتگی

مراجع

- [1] Bherwani, H., Anjum, S., Kumar, S., Gautam, S., Gupta, A., Kumbhare, H., ... & Kumar, R. (2021). "Understanding COVID-19 transmission through Bayesian probabilistic modeling and GIS-based Voronoi approach: a policy perspective," *Environment, Development and Sustainability*, 23(4), 5846-5864.
- [2] Pourghasemi, H. R., Pouyan, S., Heidari, B., Farajzadeh, Z., Shamsi, S. R. F., Babaei, S., ... & Sadeghian, F. (2020). "Spatial modeling, risk mapping, change detection, and outbreak trend analysis of coronavirus (COVID-19) in Iran (days between February 19 and June 14, 2020)," *International Journal of Infectious Diseases*, 98, 90-108.
- [3] Tabasi, M., & Alesheikh, A. A. (2017). "A Review of the Applications of Agent Based Simulation in Epidemic Diseases (Case study: Cutaneous Leishmaniasis)," *Geospatial Engineering Journal*, 8(2), 11-23, (in persian).
- [4] Franch-Pardo, I., Napoletano, B. M., Rosete-Verges, F., & Billa, L. (2020). "Spatial analysis and GIS in the study of COVID-19. A review," *Science of The Total Environment*, 739, 140033.
- [5] Giuliani, D., Dickson, M. M., Espa, G., & Santi, F. (2020). "Modelling and predicting the spatio-temporal spread of COVID-19 in Italy," *BMC infectious diseases*, 20(1), 1-10.
- [6] Pourghasemi, H. R., Pouyan, S., Farajzadeh, Z., Sadhasivam, N., Heidari, B., Babaei, S., & Tiefenbacher, J. P. (2020). "Assessment of the outbreak risk, mapping and infection behavior of COVID-19: Application of the autoregressive integrated-moving average (ARIMA) and polynomial models," *Plos one*, 15(7), e0236238.
- [7] Ramirez, I. J., & Lee, J. (2020). "COVID-19 emergence and social and health determinants in Colorado: a rapid spatial analysis," *International journal of environmental research and public health*, 17(11), 3856.
- [8] Melin, P., Monica, J. C., Sanchez, D., & Castillo, O. (2020). "Analysis of spatial spread relationships of coronavirus (COVID-19) pandemic in the world using self organizing maps," *Chaos, Solitons & Fractals*, 138, 109917.
- [9] Mollalo, A., Vahedi, B., & Rivera, K. M. (2020). "GIS-based spatial modeling of COVID-19 incidence rate in the continental United States," *Science of the total environment*, 728, 138884.
- [10] Mollalo, A., Rivera, K. M., & Vahedi, B. (2020). "Artificial neural network modeling of novel coronavirus (COVID-19) incidence rates across the continental United States," *International journal of environmental research and public health*, 17(12), 4204.
- [11] Saba, A. I., & Elsheikh, A. H. (2020). "Forecasting the prevalence of COVID-19 outbreak in Egypt using nonlinear autoregressive artificial neural networks," *Process safety and environmental protection*, 141, 1-8.
- [12] Pal, R., Sekh, A. A., Kar, S., & Prasad, D. K. (2020). "Neural network based country wise risk prediction of COVID-19," *Applied Sciences*, 10(18), 6448.
- [13] Charandabi, N. K., & Gholami, A. (2021). "COVID-19 Spatiotemporal Hotspots and Prediction Based on Wavelet and Neural Network," In *COVID-19 Pandemic, Geospatial Information, and Community Resilience* (pp. 211-226). CRC Press.
- [14] Samany, N. N., Toomanian, A., Maher, A., Hanani, K., & Zali, A. R. (2021). "The most places at risk surrounding the COVID-19 treatment hospitals in an urban environment-case study: Tehran city," *Land use policy*, 109, 105725.
- [15] Borghi, P. H., Zakordonets, O., & Teixeira, J. P. (2021). "A COVID-19 time series forecasting model based on MLP ANN," *Procedia Computer Science*, 181, 940-947.
- [16] Kafieh, R., Arian, R., Saedizadeh, N., Amini, Z., Serej, N. D., Minaee, S., ... & Haghjooy Javanmard, S. (2021). "Covid-19 in iran: Forecasting pandemic using deep learning," *Computational and mathematical methods in medicine*, 2021.
- [17] Shao, Q., Xu, Y., & Wu, H. (2021). "Spatial Prediction of COVID-19 in China Based on Machine Learning Algorithms and Geographically Weighted Regression," *Computational and Mathematical Methods in Medicine*, 2021.
- [18] Shariati, M. *et al.* (2020) "Spatial analysis of COVID-19 and exploration of its environmental and Socio-demographic risk factors using spatial statistical methods: A case study of Iran," *Health in Emergencies & Disasters Quarterly*, 5(3), pp. 145-154.
- [19] Talkhi, N. *et al.* (2021) "Modeling and forecasting number of confirmed and death caused COVID-19 in IRAN: A comparison of time series forecasting methods," *Biomedical signal processing and control*, 66(102494), p. 102494.

- [20] Mansour, S. *et al.* (2021) "Sociodemographic determinants of COVID-19 incidence rates in Oman: Geospatial modelling using multiscale geographically weighted regression (MGWR)," *Sustainable cities and society*, 65(102627), p. 102627.
- [21] Scott, L. M., & Janikas, M. V. (2010). "Spatial statistics in ArcGIS," In *Handbook of applied spatial analysis* (pp. 27-41). Springer, Berlin, Heidelberg.
- [22] Kyriakidis, P. C., & Journel, A. G. (1999). "Geostatistical space-time models: a review," *Mathematical geology*, 31(6), 651-684.
- [23] Shojaei, S. R. H., Waghei, Y., & Mohammadzadeh, M. (2018). "Geostatistical analysis of disease data: a case study of tuberculosis incidence in Iran," *Journal of Applied Statistics*, 45(8), 1476-1483.
- [24] Mitchel, A. (2005). "The ESRI Guide to GIS analysis, Volume 2: Spatial measurements and statistics," *ESRI Guide to GIS analysis*.
- [25] Saxena, R., Nagpal, B. N., Das, M. K., Srivastava, A., Gupta, S. K., Kumar, A., ... & Baraik, V. K. (2012). "A spatial statistical approach to analyze malaria situation at micro level for priority control in Ranchi district, Jharkhand," *The Indian journal of medical research*, 136(5), 776.
- [26] Zhou, H., Deng, Z., Xia, Y., & Fu, M. (2016). "A new sampling method in particle filter based on Pearson correlation coefficient," *Neurocomputing*, 216, 208-215.
- [27] Wheeler, D. C., & Páez, A. (2010). "Geographically weighted regression," In *Handbook of applied spatial analysis* (pp. 461-486). Springer, Berlin, Heidelberg.
- [28] Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2003). "Geographically weighted regression: the analysis of spatially varying relationships," *John Wiley & Sons*.
- [29] Mohammadinia, A., Saeidian, B., Pradhan, B., & Ghaemi, Z. (2019). "Prediction mapping of human leptospirosis using ANN, GWR, SVM and GLM approaches," *BMC infectious diseases*, 19(1), 1-18.
- [30] Mas, J. F., & Flores, J. J. (2008). "The application of artificial neural networks to the analysis of remotely sensed data," *International Journal of Remote Sensing*, 29(3), 617-663.
- [31] Ahmed, A. R. (2021). "Dynamics of Infectivity and Fatality of COVID-19 Pandemic," *Int J Cur Res Rev*, Vol, 13(05), 109.
- [32] Ahmadvani, M., Alesheikh, A. A., Khakifrouz, S., & Salehi-Vaziri, M. (2018). "Space-time epidemiology of Crimean-Congo hemorrhagic fever (CCHF) in Iran," *Ticks and tick-borne diseases*, 9(2), 207-216.
- [33] Lin, J. M., & Billa, L. (2021). "Spatial prediction of flood-prone areas using geographically weighted regression," *Environmental Advances*, 6, 100118.
- [34] Steel, R. G. (1960). "Principles and procedures of statistics: with special reference to the biological sciences," (No. 04; QA276, S82.).
- [35] Glantz, S. A., Slinker, B. K., & Neilands, T. B. (1990). "Primer of Applied Regression and Analysis of Variance," *McGraw-Hill. Inc., New York*.
- [36] Draper, N. R., & Smith, H. (1998). "Applied regression analysis," (Vol. 326). *John Wiley & Sons*.
- [37] Willmott, C. J., & Matsuura, K. (2005). "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Climate research*, 30(1), 79-82.
- [38] Hyndman, R. J., & Koehler, A. B. (2005). "Another look at measures of forecast accuracy," *Monash University*.
- [39] Hyndman, R. J., & Koehler, A. B. (2006). "Another look at measures of forecast accuracy," *International journal of forecasting*, 22(4), 679-688.
- [40] Pontius, R. G., Thontteh, O., & Chen, H. (2008). "Components of information for multiple resolution comparison between maps that share a real variable," *Environmental and Ecological Statistics*, 15(2), 111-142.
- [41] Willmott, C. J., & Matsuura, K. (2006). "On the use of dimensioned measures of error to evaluate the performance of spatial interpolators," *International Journal of Geographical Information Science*, 20(1), 89-102.
- [42] Tiktak, A., van der Linden, A. M., & van der Pas, L. J. (1998). "Application of the pesticide transport assessment model to a field study in a humic sandy soil in Vredepeel, The Netherlands," *Pesticide science*, 52(4), 321-336.