

Comparison of the Use of Multilayer Perceptron Neural Network with Weighted Regression in Determining the Geographical Pattern of the Brucellosis Disease Incidence in Mazandaran Province, Iran (2009-2018)

Maryam Salmani Seraji¹
Jamshid Yazdani Charati²
Reza Ali Mohammadpour Tahamtan³
Farhang Baba Mahmoudi⁴
Habib Vahedi⁵
Zahra ramazani⁶

¹ MSc Student in Biostatistics, Faculty of Health, Mazandaran University of Medical Sciences, Sari, Iran

² Professor, Department of Biostatistics and Epidemiology, Faculty of Health, Mazandaran University of Medical Sciences, Sari, Iran

³ Associate Professor, Department of Biostatistics and Epidemiology, Faculty of Health, Mazandaran University of Medical Sciences, Sari, Iran

⁴ Professor, Department of Infectious Diseases, Faculty of Medicine, Mazandaran University of Medical Sciences, Sari, Iran

⁵ Assistant Professor, Department of Basic Sciences, Faculty of Health, Mazandaran University of Medical Sciences, Sari, Iran

⁶ Postdoctoral Researcher in Biostatistics and Epidemiology, Faculty of Health, Mazandaran University of Medical Sciences, Sari, Iran

(Received September 17, 2023; Accepted October 12, 2023)

Abstract

Background and purpose: Time series models based on machine learning, including artificial neural network, perform better than classical methods. This study was performed to compare the use of a multi-layered perceptron neural network with weighted regression in determining the geographical pattern of brucellosis in Mazandaran province, Iran (2009-2018) on 3,732 patients.

Materials and methods: The study type is ecological and existing data, which is registered. Multilayer perceptron neural network was used to model and predict disease occurrence, and the results were compared with weighted regression. The root mean square error (RMSE) and mean absolute percentage error (MAPE) were used to compare the models. All analyses were performed using SPSS software (version 26) and Microsoft Excel 2016.

Results: The age-standardized rate of disease incidence was 13.2 per 100,000. The highest incidence rate (17.2) was in 2012 and the lowest (10.6) in 2014. Galugah city (35.0) and Qaemshahr and Fereydunkenar cities (3.0) had the highest and lowest incidence rates. The independent variables studied included male gender, rural residence, age over 55 years, contact with dairy products, being a homemaker, rancher and associated with livestock jobs, contact with livestock, and livestock vaccination on the standard incidence of brucellosis in both artificial neural network and weighted regression (other than variable over 55 years) models were effective.

Conclusion: The predictive power of the model in the multi-layered perceptron neural network was higher than the weighted regression.

Keywords: Brucellosis, Mazandaran, Multilayer perceptron neural network, Weighted regression

J Mazandaran Univ Med Sci 2023; 33 (Supple 2): 269-280 (Persian).

Corresponding Author: Jamshid Yazdani Charati- Faculty of Health, Mazandaran University of Medical Sciences, Sari, Iran.
(E-mail: jamshid.charati@gmail.com)

مقایسه به کارگیری شبکه عصبی پرسپترون چندلایه با رگرسیون وزن دار در تعیین الگوی جغرافیایی بروز بیماری تب مالت استان مازندران (۱۳۸۸-۱۳۹۶)

مریم سلمانی سراجی^۱
جمشید یزدانی چراتی^۲
رضاعلی محمدپور تهمتن^۳
فرهنگ بابا محمودی^۴
حبیب واحدی^۵
زهرا رضانی^۶

چکیده

سابقه و هدف: روش های یادگیری ماشین از جمله شبکه عصبی مصنوعی، پیش بینی بهتری نسبت به روش های کلاسیک دارد. این مطالعه به منظور مقایسه به کارگیری شبکه عصبی پرسپترون چندلایه با رگرسیون وزن دار در تعیین الگوی جغرافیایی بروز بیماری تب مالت استان مازندران (۱۳۸۸-۱۳۹۶) درباره ۳۷۳۲ بیمار انجام شد.

مواد و روش ها: نوع مطالعه اکولوژیکال و از نوع اگزپستینگ دیتا به صورت ثبتي است. برای مدل سازی و پیش بینی بروز بیماری از شبکه عصبی پرسپترون چندلایه استفاده و نتایج با رگرسیون وزنی مقایسه شد. جهت مقایسه مدل ها از دو معیار RMSE و MAPE استفاده شد. تمام آنالیزها با استفاده از نرم افزار SPSS v.26 و میکروسافت اکسل ۲۰۱۶ انجام شد. **یافته ها:** نرخ استاندارد سنی بروز بیماری ۱۳/۲ در ۱۰۰ هزار نفر است. سال ۱۳۹۱ بالاترین نرخ بروز (۱۷/۲) و کم ترین (۱۰/۶) مربوط به سال ۱۳۹۳ بود. شهرستان گلوگاه (۳۵/۰) و شهرستان های قائمشهر و فریدون کنار (۳/۰) بالاترین و پایین ترین نرخ بروز را داشتند. متغیرهای جنسیت مرد، سکونت روستایی، سن بالای ۵۵ سال، تماس با فراورده های لبنی، خانه دار بودن، دامداری و شغل های وابسته به دام، تماس با دام و واکسیناسیون دام بر بروز استاندارد بیماری در هر دو مدل شبکه عصبی مصنوعی و رگرسیون وزنی (به غیر از متغیر بالای ۵۵ سال) تأثیرگذار بود.

استنتاج: قدرت پیش بینی مدل در شبکه عصبی پرسپترون چندلایه از رگرسیون وزن دار بیش تر بود.

واژه های کلیدی: بروسلوز، شبکه عصبی پرسپترون چند لایه، رگرسیون وزنی، مازندران

مقدمه

پیش بینی دقیق رخداد های بیماری نیازمند داده های مدل های سری زمانی مبتنی بر یادگیری ماشین، از جمله شبکه عصبی مصنوعی، پیش بینی بهتری نسبت به روش های با کیفیت بالا و روش های آماری با کم ترین خطا هستند.

E-mail: jamshid.charati@gmail.com

مؤلف مسئول: جمشید یزدانی چراتی - ساری: کیلومتر ۱۷ جاده فرح آباد، مجتمع دانشگاهی پیامبر اعظم، دانشکده بهداشت

۱. دانشجوی کارشناسی ارشد آمار زیستی، دانشکده بهداشت، دانشگاه علوم پزشکی مازندران، ساری، ایران
 ۲. استاد، گروه آمار زیستی و اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی مازندران، ساری، ایران
 ۳. دانشیار، گروه آمار زیستی و اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی مازندران، ساری، ایران
 ۴. استاد، گروه عفونی، دانشکده پزشکی، دانشگاه علوم پزشکی مازندران، ساری، ایران
 ۵. استادیار، گروه علوم پایه، دانشکده بهداشت، دانشگاه علوم پزشکی مازندران، ساری، ایران
 ۶. دانشجوی پست دکتر آمار زیستی و اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی مازندران، ساری، ایران
- تاریخ دریافت: ۱۴۰۲/۶/۲۶ تاریخ ارجاع جهت اصلاحات: ۱۴۰۲/۸/۶ تاریخ تصویب: ۱۴۰۲/۷/۲۱

کلاسیک دارند. این روش توانمند روابط پیچیده غیرخطی بین متغیرهای مستقل و وابسته را بدون در نظر گرفتن هیچ فرضیه اولیه‌ای، مدل‌سازی می‌کند (۱) در میان روش‌های رایج شبکه عصبی مصنوعی، پرسپترون چندلایه قابلیت تعمیم‌پذیری بالاتری دارد (۲). قانون یادگیری پرسپترون چندلایه (Multilayer Perceptron): قاعده پس‌انتشار (MLP) (Error Backpropagation) خطا نام دارد که به منظور تخمین پارامترهای مجهول شبکه مورد استفاده قرار می‌گیرند. نحوه عمل پرسپترون چندلایه به این صورت است که الگویی به شبکه عرضه می‌شود و خروجی آن محاسبه می‌گردد. مقادیر خروجی واقعی و خروجی مطلوب باعث می‌گردد که ضرایب شبکه تغییر یابد؛ به گونه‌ای که در مراحل بعد، خروجی صحیح‌تری حاصل شود. برای موفقیت در آموزش شبکه باید خروجی آن را به تدریج، به خروجی مطلوب نزدیک کرد و میزان خطا را کاهش داد (۳). بروسولوز یکی از بیماری‌های عفونی شایع و مشترک بین انسان و دام است که ممکن است از طریق فراورده‌های لبنی غیرپاستوریزه، خراش‌های ناپیدای جلدی، ادرار، ترشحات دام‌های آلوده و استنشاق هوای آلوده، به انسان منتقل شود (۴). شرایط جغرافیایی ایران، مصرف مواد لبنی غیرپاستوریزه و ارتباط شغلی کارگران کشتارگاه‌ها، دامداری‌ها، دامپزشکان، قصابان و کارکنان آزمایشگاه، تب مالت را به عنوان یک بیماری شغلی مطرح می‌سازد (۵). بیماری حالت فصلی داشته و خطر بروز بیماری در فصول بهار و تابستان که زمان بارداری و زایمان دام‌ها است، بیش‌تر است (۶). این بیماری در بسیاری از کشورهای حاشیه مدیترانه، خاورمیانه، شبه‌قاره هند، آمریکای مرکزی و جنوبی به صورت اندمیک وجود دارد. به طوری که میزان بروز سالانه تب مالت در مدیترانه و کشورهای خاورمیانه ۷۸-۱ مورد در هر ۱۰۰ هزار نفر تغییر می‌کند (۷). سالانه حدود ۵۰۰ هزار نفر به تب مالت مبتلا می‌شوند (World Health Organization: WHO). ایران از جمله کشورهایی است که میزان ابتلا به تب

مالت در آن بالاست و به صورت اندمیک و در رده چهارم جهان (۸) و رتبه اول منطقه مدیترانه شرقی (۹) قرار دارد. به‌طور متوسط، در هر سال، ۲۷۵۰۰ مورد جدید بیماری گزارش شده است (۱۰). مرکز مدیریت بیماری‌های عفونی (Ministry of Health and Medical Education, Infectious Diseases Management Center) میزان بروز بیماری را در سال ۸۹، معادل ۱۵/۹ عنوان کرده است که پراکندگی آن در استان‌های کشور متفاوت است. رخداد بیماری‌های عفونی اغلب دارای یک الگوی منظم و چرخه‌ای است؛ بنابراین، تعیین دقیق و به‌هنگام تغییرات رفتار بیماری در طول زمان، امری ضروری است (۱۱). تشخیص عوامل بیماری‌زا، شناسایی تجمع مکانی بیماری و یافتن الگوی انتشار بیماری در محیط از ضروری‌ترین نیازها در زمینه بهداشت عمومی و مدیریت بیماری‌ها به شمار می‌آیند (۱۲). بنابراین، این مطالعه به منظور مقایسه به‌کارگیری شبکه عصبی پرسپترون چندلایه با رگرسیون وزن‌دار در تعیین الگوی جغرافیایی بروز بیماری تب مالت استان مازندران (۱۳۹۶-۱۳۸۸) انجام شد.

مواد و روش‌ها

جامعه پژوهش ۳۷۳۲ بیمار با تشخیص محتمل و قطعی است که با تشخیص آزمایشگاهی تحت پیگیری و درمان قرار گرفته‌اند. معیار ورود افراد ثبت دقیق و کامل اطلاعات بیماران در فرم‌های بررسی مربوطه (مورد تأیید مرکز مدیریت بیماری‌ها) بر اساس تعاریف، طبقه‌بندی شد. محتمل: مورد مظنونی که تیتراژ آزمون رایت آن دارای تیتراژ مساوی یا بیش‌تر از ۱/۸۰ باشد. قطعی: مورد مظنون یا محتملی که با معیار تشخیص‌های قطعی آزمایشگاهی همراه باشد (الف). جدا کردن عامل بیماری از نمونه‌های بالینی در محیط کشت، ب. تیتراژ (2-Mercaptoethanol:2ME) مساوی یا بیش‌تر از یک‌چهارم، ج. کومبس رایت (coombs wright) مثبت با فاصله یک دقت بالاتر از رایت (۱۳).

$$MAPE = \frac{\sum_{i=1}^n \left| \frac{O_i - P_i}{O_i} \right|}{n} * 100$$

O_i مقدار مشاهده شده،

P_i مقدار مورد انتظار و

n تعداد نمونه است.

یافته‌ها

جدول شماره ۱ بروز استاندارد سنی بیماری به تفکیک شهرستان‌ها است. در جدول شماره ۲، اطلاعات دموگرافیک و اپیدمیولوژیک بیماران گزارش شده است. بروز استاندارد سنی بیماری در زنان ۱۰/۴ و در مردان ۱۵/۶ بود. این مقدار در محیط روستایی ۲۱/۳ و در محیط شهری ۶/۴ به دست آمد. هم‌چنین، بروز استاندارد بیماری با افزایش سن، بیش‌تر شده بود، طوری که در گروه سنی ۴۵ تا ۵۴ سال با مقدار ۱۹/۴ و بالاتر از ۵۵ سال ۲۲/۶ بود. در فصول بهار و تابستان نیز به ترتیب، ۳/۹ و ۴/۴ گزارش شده بود (جدول شماره ۱). خلاصه مدل MLP در جدول شماره ۳ و آزمون رگرسیون وزن‌دار با روش گام‌به‌گام در جدول شماره ۴ ارائه شد.

جدول شماره ۱: بروز استاندارد سنی بیماری در شهرستان‌های استان مازندران (۱۳۹۶-۱۳۸۸)

نام شهر	درصد	نام شهر	درصد	نام شهر	درصد
آمل	۱۲/۸	نور	۲۵/۸	جویبار	۱۱/۱
بهشهر	۱۵/۸	نوشهر	۱۴/۰	گلرگانه	۳۵/۰
تنگابن	۱۴/۹	بابلسر	۶/۲	فریدون‌کنار	۳/۰
رامسر	۲۸/۶	بابل	۷/۷	عباس‌آباد	۸/۶
ساری	۱۴/۱	محمودآباد	۶/۹	میان‌درود	۲۰/۸
سوادکوه	۲۹/۶	نکا	۲۷/۹	سیرمغ	۷/۲
قائم‌شهر	۳/۰	چالوس	۱۲/۹		

مدل شبکه عصبی پرسپترون چندلایه

در این روش، ۱۱۶۸ رکورد داده از نوع فراوانی به سه زیرمجموعه آموزش (۷۵ درصد)، آزمون (۱۵ درصد) و ارزیابی (۱۰ درصد) تقسیم شدند. از آن جایی که لایه خروجی از یک متغیر وابسته به مقیاس تشکیل شده است، مجموع مربعات خطا در اینجا ارائه می‌شود. مجموع مربعات خطا، خطایی است که شبکه عصبی سعی می‌کند در مرحله آموزش، آن را به حداقل برساند.

در این مطالعه، پس از گرفتن مجوزهای مربوطه و هماهنگی بین دانشگاه علوم پزشکی مازندران و مرکز مدیریت بیماری‌های واگیر، اطلاعات مورد نیاز در غالب نرم‌افزار اکسل دریافت شد. نوع مطالعه اکولوژیکال و از نوع اگزپستینگ دیتا و به صورت ثبتي است. برای مدل‌سازی و پیش‌بینی بروز بیماری از شبکه عصبی پرسپترون چندلایه استفاده و نتایج با مدل رگرسیون وزنی مقایسه شد. در برازش مدل، شبکه عصبی سه‌لایه متشکل از لایه ورودی شامل متغیرهای مرد بودن، روستایی بودن، سن بالای ۵۵ سال، شغل خانه‌دار، دامداری و شغل‌های مربوط به دام، تماس با دام، تماس با فراورده‌های لبنی و سابقه واکسیناسیون دام، یک لایه پنهان و لایه خروجی بروز استاندارد سنی در دو فاکتور شهرستان و سال برای ۱۱۶۸ رکورد داده در نظر گرفته شد. تانژانت هایپربولیک به عنوان تابع فعالیت لایه خروجی و لایه پنهان در نظر گرفته شد. روش مقیاس‌بندی مجدد متغیرهای کمکی استاندارد شده و متغیر وابسته به مقیاس نرمال تنظیم شده تعیین شد. در روش رگرسیون وزنی، جمعیت شهرستان‌ها به عنوان متغیر وزن در نظر گرفته شد. برای محاسبه بروز از داده‌های سرشماری جمعیت (Statistical Center of Iran: SCI) و نرخ رشد جمعیت سال‌های مورد مطالعه و سپس، برای محاسبه بروز استاندارد سنی (Age-Standardized Rate: ASR)، از روش استانداردسازی مستقیم با استفاده از جمعیت استاندارد WHO استفاده شد (۱۴). مقادیر گم‌شده با استفاده از روش MICE (Multivariate imputation by chained equations) برآورد شد. جهت مقایسه مدل‌ها از دو معیار RMSE (Root mean square error) و MAPE (Mean absolute percentage error) استفاده شد. تمام آنالیزها با استفاده از نرم‌افزار SPSS v.26 و مایکروسافت اکسل ۲۰۱۶ انجام شد.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (O_i - P_i)^2}{n}}$$

جدول شماره ۳: خلاصه مدل شبکه عصبی MLP

گروه	مجموع مربعات خطا	خطای نسبی
آموزش	۲,۹۲۸	۰/۰۹۷
آزمون	۰,۶۶۳	۰/۱۴۵
ارزیابی	-	۰/۱۴۹

قانون توقف استفاده شده یک مرحله متوالی بدون کاهش خطا است. خطاهای نسبی برای آموزش، آزمون و ارزیابی بسیار مشابه هستند (به ترتیب، ۰/۰۹۷، ۰/۱۴۵ و ۰/۱۴۹)، از این رو می توان فرض کرد که مدل بیش از حد آموزش دیده نیست. هم چنین، مقادیر پایین، نشان دهنده مطلوبیت بالای مدل است. محاسبه خطاها بر اساس مجموعه آزمون بود. در نمودار پراکنش مقادیر مشاهده شده و برآزش شده، هرچه نقاط به نیمساز اول و سوم نزدیک تر باشند و هم چنین، ساختار نامشخصی بین مقادیر برآزش شده و باقی مانده وجود داشته باشد، قدرت پیش بینی بالاتر است (مقادیر بین خطوط ۲، ۲-) متمرکز شده اند) (نمودار شماره ۱ و ۲).

تحلیل اهمیت متغیرهای مستقل

اهمیت متغیرهای مستقل در تشخیص این نکته است که به چه میزان مقادیر پیش بینی شده توسط شبکه، با تغییر مقادیر متغیر مستقل، تغییر می نماید. با ورود متغیرهای شهرستان و سال، متغیرهای مرد بودن (۱۰۰ درصد)، شهرستان (۷۸/۷ درصد)، خانه دار بودن (۷۳/۴ درصد) و تماس با فرآورده های لبنی (۲/ درصد) بالاترین اهمیت و متغیرهای سال (۱۰ درصد) و دامداری و شغل های مربوط به دام (۱۰/۷ درصد) کم ترین اهمیت را در پیش بینی بروز استاندارد بیماری داشت (نمودار شماره ۳).

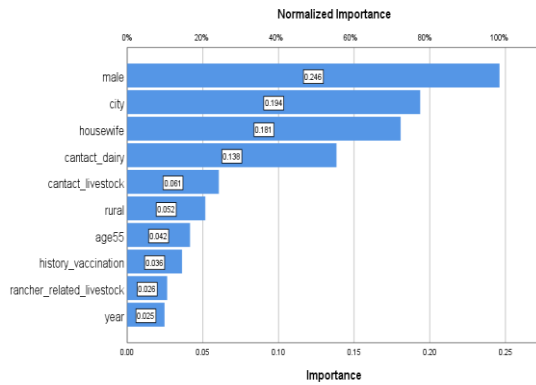
مدل رگرسیون وزن دار

با در نظر گرفتن جمعیت به عنوان متغیر وزن، رگرسیون وزن دار در سطوح فاکتور شهرستان و سال اجرا شد. تمام متغیرهای مستقل به غیر از سن بالای ۵۵ سال، بر بروز استاندارد بیماری تب مالت تأثیر گذار است.

جدول شماره ۲: توزیع فراوانی (درصد) اطلاعات دموگرافیک و

جنسیت	تعداد (درصد)
زن	۳۹(۱) ۱۴۶۰
مرد	(۶/۹) ۲۲۷۲
محل سکونت	
شهری	(۲۷/۴) ۱۰۲۳
روستایی	(۷۲/۶) ۲۷۰۹
شغل	
کودک	(۲/۶) ۹۷
محصل	(۸/۹) ۳۳۳
خانه دار	(۳۰) ۱۱۱۸
کشاورز	(۲/۹) ۱۰۹
دامدار	(۱۴/۹) ۵۵۶
مرتبط با دام	(۱۴/۴) ۵۳۷
سایر	(۲۴/۸) ۹۲۴
پاسخ نداده	(۱/۶) ۵۸
گروه سنی	
۱۴-۰	(۸/۱) ۳۰۲
۱۵-۲۴	(۱۴/۲) ۵۳۰
۲۵-۴۴	(۳۷/۲) ۱۳۹۰
۴۵-۵۴	(۱۶/۹) ۶۳۰
۵۵=>	(۲۳/۳) ۸۷۰
پاسخ نداده	(۰/۳) ۱۰
فصل	
بهار	(۲۹/۷) ۱۱۰۸
تابستان	(۳۳/۶) ۱۲۵۳
پاییز	(۲۰) ۷۴۷
زمستان	(۱۶/۷) ۶۲۴
ملیت	
ایرانی	(۹۹/۴) ۳۷۱۰
غیر ایرانی	(۰/۶) ۲۲
سابقه واکسیناسیون دام	
دارد	(۳۳/۸) ۱۲۶۳
ندارد	(۲۴/۸) ۹۲۴
پاسخ نداده	(۴۱/۴) ۱۵۴۵
سابقه تماس با دام	
دارد	(۵۹/۵) ۲۲۲۱
ندارد	(۳۷/۱) ۱۳۸۴
پاسخ نداده	(۳/۴) ۱۲۷
نوع تماس با دام	
تماس با دام زنده	(۵۵/۱) ۲۰۵۶
ذبح دام، تماس با خون و ترشحات حضور در محل نگهداری دام و کود حیوانی	(۱۸/۱) ۶۷۶
ابتلای سایر افراد خانواده	(۹/۸) ۳۶۵
دارد	(۱۳/۱) ۴۹۰
ندارد	(۷۱/۲) ۲۶۵۷
پاسخ نداده	(۱۵/۷) ۵۸۵
سابقه استفاده از فرآورده لبنی	
دارد	(۸۰) ۲۹۸۴
ندارد	(۱۰/۳) ۳۸۴
پاسخ نداده	(۹/۸) ۳۶۴
نوع فرآورده لبنی	
شیر	(۴۴/۶) ۱۶۶۴
پنیر	(۴۷/۴) ۱۷۷۰
خامه	(۴/۲) ۱۵۵
کره	(۸/۲) ۳۰۷
آغوز	(۲/۶) ۹۶
سرشیر	(۷/۹) ۲۹۶
بستی	(۳/۱) ۱۱۶

محصل: دانش آموز و دانشجو. دامدار: دامدار، کارگر دامداری. مرتبط با دام: قصاب، کارگر کشتارگاه، کارمند- دامدار، کشاورز- دامدار. سایر: کارگر، کارمند، آزاد، راننده، نظامی، بیکار، کارمند- کشاورز. ذبح دام، تماس با خون و ترشحات: تماس با ترشحات زایمانی یا جنین سقط شده دام، تماس با لاشه و ترشحات پس از ذبح، تماس با گوشت و اعضای بدن. ثبت آیتم حضور در محل نگهداری دام و کود حیوانی از سال ۹۳ (۱۵۶۱ مورد).



نمودار شماره ۳: اهمیت نرمال شده متغیرهای ورودی به شبکه عصبی پرسپترون چندلایه در تعیین الگوی جغرافیایی بروز بروسلوز در استان مازندران از ۱۳۸۸-۱۳۹۶

متغیرهای مرد بودن، خانه‌دار بودن، تماس با فراورده‌های لبنی، روستایی بودن و شهرستان بیش‌ترین اهمیت و متغیرهای دامداری و شغل‌های مربوط به دام، سابقه واکسیناسیون دام، تماس با دام و سال کم‌ترین اهمیت را در بروز استاندارد بیماری داشتند. مدلی با کم‌ترین خطای استاندارد، بیش‌ترین ضریب تعیین تنظیم شده و بیش‌ترین ضریب تعیین پیش‌بینی شده، از قدرت پیش‌بینی بالاتری برخوردار است.

در نمودار احتمال نرمال، نزدیکی نقاط به خط بیانگر نرمال بودن باقی‌مانده‌ها است. در نمودار مقادیر پیش‌بینی شده و باقی‌مانده، نبود ساختار مشخص بیانگر ثبات واریانس و پیش‌بینی مطلوب مدل است. در نمودار هیستوگرام باقی‌مانده‌ها نیز هرچه چیدمان مستطیل‌ها به منحنی نرمال نزدیک‌تر باشد، بیانگر نرمال بودن باقی‌مانده‌ها است. در نمودار باقی‌مانده در برابر ترتیب جمع‌آوری داده، نبود روند مشخص در مدل نشان‌دهنده

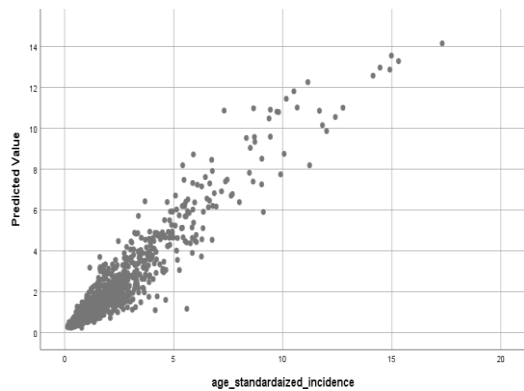
مستقل بودن باقی‌مانده‌ها است (نمودار شماره ۴). هر ۴ آیت‌مدنظر در نمودارهای بالا دیده می‌شود و نشان‌دهنده مناسب بودن قدرت پیش‌بینی مدل است. جدول شماره ۴ نشان می‌دهد که مدل پرسپترون چندلایه به‌طور کلی، بهتر از مدل رگرسیونی با مقادیر کم‌تر RMSE و MAPE عمل می‌کند.

جدول شماره ۴: به‌کارگیری آزمون رگرسیون وزن‌دار گام‌به‌گام تعیین الگوی جغرافیایی بروز بروسلوز و عوامل مرتبط با آن در استان مازندران طی سال‌های ۱۳۸۸-۱۳۹۶

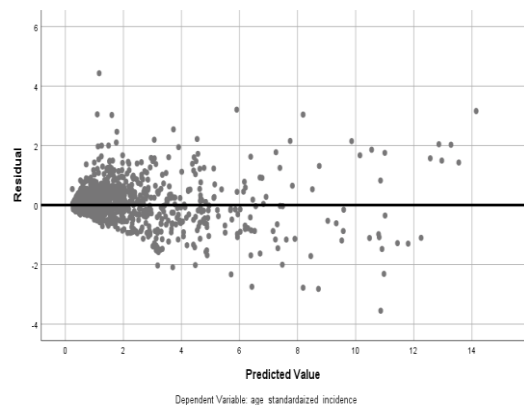
متغیرهای مستقل	ضریب رگرسیونی	سطح معنی‌داری	پارامترهای انتخاب بهترین مدل
مقدار ثابت	۲/۹۸۷	۰/۰۰۱	R-sq(pred) ۹۷/۶۷ درصد
روستایی بودن	۰/۴۶۵	۰/۰۰۱	R-sq(adj) ۹۷/۸۶ درصد
تماس با فراورده‌های لبنی	۰/۵۸۴	۰/۰۰۱	R-sq ۹۷/۹۳ درصد
مرد بودن	۱/۱۸۶	۰/۰۰۱	
شغل خانه‌دار	۰/۹۱۸	۰/۰۰۱	
شهرستان	۰/۴۹۳	۰/۰۰۱	
سال	۰/۲۸۳	۰/۰۰۱	S ۲۴۸/۱۱۴ درصد
سابقه واکسیناسیون دام	۰/۱۲۷	۰/۰۰۱	
تماس با دام	۰/۱۳۷	۰/۰۰۱	
دامداری و شغل‌های مربوط به دام	۰/۰۰۵	۰/۰۰۲	

جدول شماره ۵: مقایسه RMSE و MAPE برای هر دو مدل رگرسیون وزن‌دار و شبکه‌های مصنوعی

خطاها	شبکه عصبی پرسپترون	رگرسیون وزنی
RMSE	۱/۸۸۴	۲/۹۶۰
MAPE	۱۰/۷۲۳۸	۱۳۶/۸۳۶

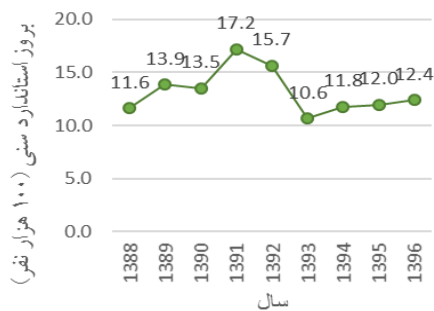


نمودار شماره ۱: پراکنش مقادیر برآزش شده در شبکه عصبی پرسپترون چندلایه در مقابل مقادیر مشاهده شده در تعیین الگوی جغرافیایی بروز بروسلوز در استان مازندران از ۱۳۸۸-۱۳۹۶



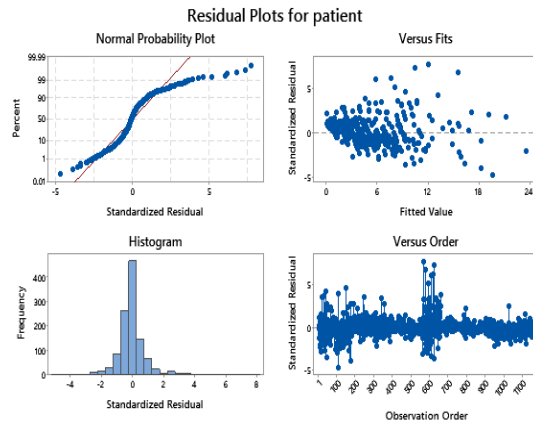
نمودار شماره ۲: پراکنش مقادیر برآزش شده در شبکه عصبی پرسپترون چندلایه در مقابل مقادیر باقیمانده در تعیین الگوی جغرافیایی بروز بروسلوز در استان مازندران از ۱۳۸۸-۱۳۹۶

شهرستان‌ها، گلوگاه (۳۵/۰) بروز بیش‌تر و قائمشهر (۳/۰) و فریدون‌کنار (۳/۰) بروز کم‌تری داشتند. این اختلاف جغرافیایی می‌تواند به دلیل نسبت متفاوت جمعیت شهری و روستایی و هم‌چنین، حساسیت متفاوت شهرستان‌ها در خصوص نظام مراقبت و گزارش‌دهی باشد. از طرفی، با وجود تعداد دام بیش‌تر در برخی شهرستان‌ها و آمیخته بودن شغل افراد با نگاه‌داری دام، این اختلاف مورد انتظار است (۱۸).



نمودار شماره ۵: بروز استاندارد سنی بیماری تب مالت استان مازندران (۱۳۸۸-۱۳۹۶)

قدرت پیش‌بینی مدل در شبکه عصبی پرسپترون از رگرسیون وزن‌دار بیش‌تر بود (مقادیر کم‌تر RMSE و MAPE). با ورود متغیرهای شهرستان و سال، متغیرهای مرد بودن، شهرستان، شهرستان، خانه‌دار بودن و تماس با فرآورده‌های لبنی به ترتیب، بالاترین اهمیت و متغیرهای سال و دامداری و شغل‌های مربوط به دام، کم‌ترین اهمیت را در پیش‌بینی بروز استاندارد بیماری داشتند. در رگرسیون وزنی نیز متغیرهای مرد بودن، خانه‌دار بودن، تماس با فرآورده‌های لبنی، روستایی بودن و شهرستان به ترتیب، بیش‌ترین اهمیت و متغیرهای دامداری و شغل‌های مربوط به دام، سابقه واکسیناسیون دام، تماس با دام و سال کم‌ترین اهمیت را در بروز استاندارد بیماری داشتند. در رگرسیون وزنی، متغیر سن بالای ۵۵ سال از مدل حذف شده بود. دلیل پیش‌بینی بالای مدل وجود عوامل خطر بیماری تب مالت در این پژوهش بوده است. متغیرهای دموگرافیک شامل مرد بودن، روستایی بودن، خانه‌دار بودن و دامدار بودن و متغیرهای



نمودار شماره ۴: نمودار باقیمانده های ناشی از بکارگیری رگرسیون وزن دار در تعیین الگوی جغرافیایی بروز بروسوز در استان مازندران از سال ۱۳۸۸-۱۳۹۶

بحث

روند بروز بیماری در گذر سال‌ها متغیر بوده است. این روند تا سال ۹۱ با افزایش سریع بروز و پس از آن تا سال ۹۳، با کاهش سریع و در سال‌های اخیر، با افزایش تدریجی همراه بود (نمودار شماره ۵). دلیل این امر را به ترتیب تغییرات روند، می‌توان بهبود نظام آماری و ثبت اطلاعات در سیستم اتوماسیون بیماری‌ها، افزایش بهداشت و واکسیناسیون دام‌ها و تمایل بیش‌تر مردم به استفاده از فرآورده‌های لبنی سنتی و واردات بی‌رویه دام‌ها از همسایگان بومی مانند عراق، پاکستان و افغانستان دانست. به دلیل نبود برنامه‌های کنترل دقیق برای بیماری‌های حیوانات در این کشورها، تب مالت در این مناطق بومی است و خطر انتقال تب مالت از این کشورها به ایران وجود دارد (۱۵). طبق گزارش وزارت بهداشت، استان مازندران دارای آلودگی متوسط در سال ۸۹ بود. در مطالعه حاضر، این مقدار در طی ۹ سال ۱۳/۲ بود که مطابق با آمار وزارت بهداشت است و نشان می‌دهد که مازندران هم‌چنان با آلودگی متوسط مواجه است. در یک مطالعه کوهپور مبتنی بر جمعیت، میزان بروز بیماری در مازندران، ۹/۴۰ تا ۲۰/۰۸ گزارش شده بود (۱۶). یک مطالعه همه‌گیرشناسی مکانی-زمانی تب مالت در استان‌های شمالی نشان داد که بروز بیماری در سال‌های اخیر افزایش داشته است (۱۷). در میان

اپیدمیولوژیک شامل تماس با دام و تماس با فراورده‌های لبنی، همگی از مهم‌ترین عوامل خطر این بیماری محسوب می‌شوند. عملکرد شبکه عصبی روابط غیرخطی میان پدیده‌ها را در نظر می‌گیرد و یکی از کاربردهای شبکه عصبی دسته‌بندی و شناسایی الگو است؛ بنابراین، تفکیک به متغیرهای طبقه‌ای، باعث تکرار بیش‌تر الگو می‌شود و مدل را در رسیدن به یک ساختار مناسب راهنمایی می‌کند، در حالی که عملکرد رگرسیون وزنی تنها خطی است. دلیل متفاوت بودن شاخص ارزیابی می‌تواند انتخاب یک متغیر وزنی مناسب با توجه به شرایط متغیرهای موجود در مدل باشد.

در بیش‌تر مطالعات پیشین، مناسب‌تر بودن عملکرد شبکه عصبی از جمله پرسپترون، نسبت به سایر مدل‌ها مشهود بود که با نتایج پژوهش حاضر هماهنگ است. اولیویرا و همکاران (۲۰۱۷) در پژوهش داده‌کاوی خود به شناسایی عوامل اصلی مؤثر در تأخیر گزارش موارد HIV-AIDS در سیستم نظارتی پرتغال پرداختند. روش‌های مورد استفاده شامل شبکه‌های عصبی مصنوعی پرسپترون چندلایه، طبقه‌بندی کننده بیزی ساده، ماشین‌های بردار پشتیبان و الگوریتم نزدیک‌ترین همسایه K است. بالاترین دقت طبقه‌بندی صحت و فراخوانی برای MLP به دست آمد (۱۹).

در مطالعه جیوو وانگ و همکاران (۲۰۱۷)، یک رویکرد شبکه عصبی مصنوعی مبتنی بر داده‌سنجش از دور برای پیش‌بینی بروز بیماری‌های عفونی حساس به آب‌وهوا (تب مالت انسانی) مورد پژوهش قرار گرفت. مدل شبکه عصبی مصنوعی جدیدی با استفاده از تعداد موارد سالانه HB در سطح شهرستان و با ۳۷ متغیر محیطی توسعه داده شد. شاخص پوشش گیاهی پیشرفته به عنوان مهم‌ترین پیش‌بینی کننده بروز شناخته شد و دمای سطح زمین و سایر متغیرهای مربوط به دما و بارندگی در رتبه‌های بعدی بودند. این مطالعه نشان می‌دهد که با استفاده از مدل شبکه عصبی مصنوعی و متغیرهای محیطی، بروز HB ممکن است با صحت

معقولی پیش‌بینی شود (۲۰). تاپاک و همکاران (۲۰۱۹)، در پژوهش خود به ارزیابی مقایسه‌ای مدل‌های سری زمانی برای پیش‌بینی شیوع آنفلوانزا در ایران پرداختند. مطالعه حاضر به بررسی صحت ماشین بردار پشتیبان، مدل‌های سری زمانی شبکه عصبی مصنوعی و جنگل تصادفی ILI و تشخیص بروز پرداخته بود. مدل سری زمانی جنگل تصادفی از سه روش دیگر در مدل‌سازی فرکانس‌های ILI هفتگی برای مجموعه آزمون بهتر بود. همچنین، شبکه عصبی در تشخیص بروز با دقت کل ۰/۸۸۹ برای مجموعه آزمایش بهتر بود. نتایج نشان داد که مدل‌های سری زمانی می‌توانند برای پیش‌بینی فراوانی ILI و بروز هفتگی به طور مؤثر مورد استفاده قرار گیرند (۲۱).

جی وانگ و همکاران (۲۰۱۹) در مطالعه خود، با استفاده از یک شبکه عصبی حافظه کوتاه‌مدت طولانی (یک روش افزایش یادگیری عمیق)، به پیش‌بینی بروز HIV در گوانگشی چین در کنار مدل‌های میانگین متحرک جمع بسته اتورگرسیون، مدل‌های شبکه عصبی رگرسیون کلی و هموارساز نمایی پرداختند. مدل LSTM نسبت به سایر مدل‌های سری زمانی مؤثرتر بود و برای نظارت و کنترل بیماری‌های اپیدمی محلی HIV مهم شناخته شد (۲۲). آهنگرکانی و دیگران (۱۳۹۵)، در پژوهش خود به منظور تهیه نقشه پیش‌بینی توزیع مکانی-زمانی بیماری لپتوسپیروز، از دو مدل رگرسیون وزن‌دار جغرافیایی و شبکه عصبی پرسپترون چندلایه استفاده کردند. نتایج هر دو مدل نشان می‌دهند که شبکه عصبی MLP در حالت دوم (داده‌های مربوط به چند سال)، توانسته است نرخ شیوع بیماری را بهتر از مدل رگرسیون وزن‌دار جغرافیایی و حالت اول (داده‌های مربوط به یک سال) آموزش شبکه عصبی پیش‌بینی کند. علت این نتیجه را می‌توان این گونه بیان نمود که در حالت دوم، به دلیل اینکه از داده‌های مربوط به سال‌های بیش‌تری استفاده شده است، شبکه قادر بوده با دقت بیش‌تری آموزش ببیند و در نتیجه، خروجی

دقیق‌تری حاصل نماید (۲۳). تاپاک و همکاران (۱۳۹۷)، در پژوهش خود به ارزیابی و مقایسه عملکرد سه روش یادگیری ماشین در مدل‌سازی و پیش‌بینی سری زمانی بروسولوز بر اساس پارامترهای اقلیمی پرداختند. مدل شبکه‌ی عصبی چندلایه پرسپترون در پیش‌بینی داده‌های مورد مطالعه، عملکرد بهتری داشت. دما نسبت به سایر پارامترهای اقلیمی، عامل مؤثرتری در بروز این بیماری بود. بنابراین، شبکه عصبی چندلایه پرسپترون می‌تواند به‌عنوان یک روش کارا برای تشخیص رفتار روند بروسولوز در طول زمان به کار رود (۲۴).

فرامرزی و همکاران (۱۳۹۸)، در مطالعه خود به اثرات بالقوه پارامترهای اقلیمی بر بروز تب مالت انسان در استان فارس (۲۰۰۹-۲۰۱۵) پرداختند. با استفاده از مدل شبکه عصبی مصنوعی (چهار لایه، چهار نورون در هر لایه) الگوریتم آموزش لونیگ-مارکوارت و تابع انتقال سیگموئید مشخص شد که میزان همبستگی ۰/۸۹ در سطح تمرین و ۰/۸۰ در سطح آزمون (با کم‌ترین میزان خطا)، بهترین مقادیر در مدل‌سازی پرسپترون چندلایه است (۲۵). ملالو و همکاران (۲۰۱۹) از یک شبکه عصبی مبتنی بر GIS برای توزیع مکانی سل در سرتاسر ایالات متحده استفاده کردند. در بین مدل‌های توسعه‌یافته، MLP تک‌لایه پنهان بهترین دقت آزمایش را داشت. تجزیه و تحلیل حساسیت مدل MLP نشان داد که نسبت جمعیت مهاجر، اقشار تحت نظارت جمعیت و حداقل دما دارای بیش‌ترین سهم بودند. یافته‌های این مطالعه می‌تواند بینش مفیدی برای اولویت‌بندی تخصیص منابع در مناطق مستعد خطر ایجاد کند (۲۶). در برخی از مطالعات نیز مدل‌های دیگر عملکرد بالاتری نسبت به شبکه عصبی داشتند. ژانگ و همکاران (۲۰۱۴) در مطالعه خود، از ۹ نوع داده بیماری عفونی برای ارزیابی و مقایسه عملکرد چهار روش سری زمانی، یعنی دو روش تجزیه (رگرسیون و هموارساز نمایی)، میانگین متحرک جمع بسته اتورگرسیون و ماشین بردار پشتیبانی استفاده کردند. این مقایسه نشان داد که

هیچ روش واحدی کاملاً برتر از دیگران نیست، با این حال SVMها در بیش‌تر موارد، از مدل ARIMA و روش‌های تجزیه بهتر عمل می‌کنند (۲۷). محمدی نیا و همکاران (۲۰۱۹) در پژوهش خود به پیش‌بینی نقشه‌بندی لپتوسپیروز انسانی استان گیلان با استفاده از چهار رویکرد توانمند پرداختند. کارایی مدل‌ها به شرح زیر بود: $GWR > SVM > GLM > ANN$. رویکردهای به کاررفته می‌توانند لپتوسپیروز را با سطح اهمیت بالایی مدل‌سازی و پیش‌بینی کنند. دما و رطوبت مؤثرترین پارامترهای مورد بررسی شناخته شدند (۲۸). محمدی نیا و همکاران (۱۳۹۴) در مطالعه خود به مقایسه رگرسیون وزن‌دار جغرافیایی، شبکه عصبی مصنوعی و الگوریتم ژنتیک در مدل‌سازی مکانی تب شالیزار و تأثیر عوامل محیطی بر آن (۲۰۰۹-۲۰۱۱) در سطح دهستان پرداختند. نتایج حاکی از آن است که به ترتیب، روش رگرسیون وزن‌دار جغرافیایی، شبکه عصبی مصنوعی و الگوریتم ژنتیک نتایج بهتری ارائه نمودند. از طرفی، با ورود همه سال‌ها به طور هم‌زمان به هریک از روش‌ها، کاهش ضریب تبیین شبکه عصبی مصنوعی و افزایش نسبی آن در الگوریتم ژنتیک و رگرسیون وزن‌دار جغرافیایی دیده شد. هم‌چنین، متغیرهای متوسط دما، رطوبت و تبخیر بیش‌ترین تأثیر را بر بیماری تب شالیزار داشتند (۲۹).

باقری و همکاران (۱۳۹۸) در یک مطالعه کوهورت مبتنی بر جمعیت به پیش‌بینی میزان بروز ماهانه تب مالت با مدل‌های داده‌کاوی ماشین بردار پشتیبانی، MARS و جنگل تصادفی، علاوه بر یک مدل کلاسیک میانگین متحرک جمع بسته اتورگرسیون پرداختند. مدل MARS نسبت به مدل‌های دیگر برای پیش‌بینی مناسب‌تر بود (۱۶). برای رسیدن به دقت پیش‌بینی بالاتر، می‌توان از ترکیبی از مدل‌های رگرسیون وزنی و شبکه عصبی مصنوعی، از جمله رگرسیون وزنی شبکه عصبی از نظر جغرافیایی با داشتن داده‌های تعریف شده برای این مدل‌ها، بهره برد.

بروز استاندارد بیماری در هر دو مدل شبکه عصبی مصنوعی و رگرسیون وزنی (به غیر از متغیر بالای ۵۵ سال) تأثیر گذار بود. قدرت پیش‌بینی مدل در شبکه عصبی پرسپترون از رگرسیون وزن‌دار بیش تر بود.

سپاسگزاری

این مطالعه حاصل طرح تحقیقاتی مصوب معاونت تحقیقات و فناوری دانشگاه علوم پزشکی مازندران با کد اخلاق IR.MAZUMS.REC.1398.3607 و قسمتی از پایان نامه مریم سلمانی سراجی، دانشجوی کارشناسی ارشد آمار زیستی، است. از کارشناسان محترم حوزه معاونت بهداشتی دانشگاه علوم پزشکی مازندران به خاطر مساعدت در جمع‌آوری اطلاعات تشکر می‌گردد.

در مطالعات بسیاری، به ویژه مطالعات جغرافیایی، از مدل‌های رگرسیون وزن‌دار جغرافیایی و شبکه عصبی مصنوعی تنها با حضور متغیرهای محیطی و توپوگرافی استفاده شد. در این مطالعه، اطلاعات بروز بیماری را به همراه ریسک فاکتورهای احتمالی به اطلاعات جغرافیایی مورد نظر اضافه نمودیم که ترکیبی از متغیرهای دموگرافیک و اپیدمیولوژیک، متغیرهای مربوط به گزارش تعداد دام و تعداد موارد مثبت بروسلوز است.

در پایان، می‌توان نتیجه گرفت که متغیرهای مستقل شامل جنسیت مرد، سکونت روستایی، سن بالای ۵۵ سال، تماس با فراورده‌های لبنی، خانه‌دار بودن، دامداری و شغل‌های مربوطه، تماس با دام و واکسیناسیون دام بر

References

1. Baboo SS, Shereef IK. An efficient weather forecasting system using artificial neural network. *Int J Environ Sci Dev* 2010; 1(4): 321-326.
2. Amid S, Mesri Gundoshmian T, Shahgoli G. Comparison of MLP and RBF neural networks performance for estimation of broiler output energy. *Iranian Journal of Biosystems Engineering*. 2016; 47(2): 319-128 (Persian).
3. Hu X, Weng Q. Estimating impervious surfaces from medium spatial resolution imagery using the self-organizing map and multi-layer perceptron neural networks. *Remote Sensing of Environment* 2009; 113(10): 2089-2102.
4. Chin J. Control of communicable diseases manual. Washington: American Public Health Association 2000.
5. Shakurnia A, Ghasemzade A, Afra M, Sarizadeh G, Javidan S, Khodadadi A, et al. The Seroprevalence of brucellosis among blood donors referred to blood centers of Khuzestan province. *Sci J Iran Blood Transfus Organ* 2014; 11(3): 180-189 (Persian).
6. Baddour M. Diagnosis of brucellosis in humans: a review. *J Vet Adv* 2012; 2(4): 149-156.
7. Kaoud H ZM E-DA, Nasr SA epidemiology of brucellosis among farm animals. *Nature and Science* 2010; 5(5): 190-197.
8. Haran M AA, Kupfer Y, Seneviratne C, Chawla K, Tessler S. Brucellosis presenting as septic shock. *BMJ Case Rep* 2011; 2011: bcr1220103586.
9. Roya N, Abbas B. Colorectal cancer trends in Kerman province, the largest province in Iran, with forecasting until 2016. *Asian Pac J Cancer Prev* 2013; 14(2): 791-793.
10. Mostafavi E, Asmand M. Trend of brucellosis in Iran from 1991 to 2008. *Iranian Journal of Epidemiology* 2012; 8(1): 93-100 (Persian).

11. Box GE, Jenkins GM, Reinsel GC, Ljung GM. Time series analysis: forecasting and control. John Wiley & Sons 2015.
12. Akhavan P, Karimi M, Pahlavani P. Risk mapping of cutaneous leishmaniasis via a fuzzy C means-based neuro-fuzzy inference system. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 2014; 40: 19-23.
13. Norouzinezhad F, Erfani H, Norouzinejad A, Kaveh F, Ghaffari F. Epidemiological indices and trend of incidence of human brucellosis in Khorasan-e-Razavi Province from 2009 to 2016. Journal of Military Medicine 2019; 21(4): 362-371 (Persian).
14. Ahmad OB, Boschi-Pinto C, Lopez AD, Murray CJ, Lozano R, Inoue M. Age standardization of rates: a new WHO standard. Geneva: World Health Organization 2001; 9(10): 1-14.
15. Esmaeili H. Brucellosis in Islamic republic of Iran. J Med Bacteriol 2014; 3(3-4): 47-57 (Persian).
16. Bagheri H, Tapak L, Karami M, Amiri B, Cherghi Z. Epidemiological features of human brucellosis in Iran (2011-2018) and prediction of brucellosis with data-mining models. J Res Health Sci 2019; 19(4): e00462.
17. Seyedalizadeh N, Alesheikh A, Ahmadkhani M. Spatio-temporal and Ecological Analysis of Brucellosis in North of Iran. Journal of Geomatics Science and Technology 2019; 8(4): 221-231(Persian).
18. Mohammadian M, Salehiniya H, Kazaei S, Ramazanpour J, Mohammadian-Hafshejani A. Epidemiological characteristics and incidence rate of brucellosis in Isfahan province, Iran, 2012. Journal of Isfahan Medical School 2015; 33(355): 75-82 (Persian).
19. Oliveira A, Faria BM, Gaio AR, Reis LP. Data mining in HIV-AIDS surveillance system. J Med Syst 2017; 41(4): 51.
20. Wang J, Jia P, Cuadros DF, Xu M, Wang X, Guo W, et al. A remote sensing data based artificial neural network approach for predicting climate-sensitive infectious disease outbreaks: a case study of human brucellosis. Remote sensing 2017; 9(10): 1018.
21. Tapak L, Hamidi O, Fathian M, Karami M. Comparative evaluation of time series models for predicting influenza outbreaks: Application of influenza-like illness data from sentinel sites of healthcare centers in Iran. BMC Res Notes 2019; 12(1): 353.
22. Wang G, Wei W, Jiang J, Ning C, Chen H, Huang J, et al. Application of a long short-term memory neural network: a burgeoning method of deep learning in forecasting HIV incidence in Guangxi, China. Epidemiol Infect 2019; 147: e194.
23. Ahangarcani M, Farnaghi M, Shirzadi MR. Predictive Map of Spatio-Temporal Distribution of Leptospirosis Using Geographical Weighted Regression and Multilayer Perceptron Neural Network Methods. Journal of Geomatics Science and Technology 2016; 6(2): 79-98 (Persian).
24. Tapak L, Shirmohammadi-Khorram N, Hamidi O, Maryanaji Z. Predicting the frequency of human brucellosis using climatic indices by three data mining techniques of radial basis function, multilayer perceptron and nearest Neighbor: A comparative study. Iranian Journal of Epidemiology 2018; 14(2): 153-165 (Persian).
25. Faramarzi H, Nasiri M, Khosravi M, Keshavarzi A, Ardakani ARR. Potential effects of climatic parameters on human brucellosis in Fars province, Iran, during 2009-2015. Iran J Med Sci 2019; 44(6): 465.
26. Mollalo A, Mao L, Rashidi P, Glass GE. A GIS-based artificial neural network model for

- spatial distribution of tuberculosis across the continental United States. *Int J Environ Res Public Health* 2019; 16(1): 157.
27. Zhang X, Zhang T, Young AA, Li X. Applications and comparisons of four time series models in epidemiological surveillance data. *PLoS One* 2014; 9(2): e88075.
28. Mohammadinia A, Saeidian B, Pradhan B, Ghaemi Z. Prediction mapping of human leptospirosis using ANN, GWR, SVM and GLM approaches. *BMC infectious diseases* 2019; 19(1): 1-18.
29. Ali Mohammadi A MA. Comparison of Geographicaly Weighted Regression, Artificial Neural Network and Genetic Algorithm in Spatial Modeling of Paddy Fever. Khaje Nasir al-Din Tusi University of Technology-School of Surveying 2016.