



مدیریت آب و آبیاری

(نشریه علمی)

دوره ۱۱ ■ شماره ۳ ■ پاییز ۱۴۰۰

صفحه‌های ۴۰۹-۴۱۹

DOI: 10.22059/jwim.2021.323839.871

مقاله پژوهشی:

مقایسه مدل‌های ماشین بردار و شبکه عصبی تابع شعاعی در پیش‌بینی کیفیت آب سیمینه‌رود

بهاره حسین‌پناهی^۱، سامان نیک‌مهر^۲، کیومرث ابراهیمی^{۳*}

۱. دانشجوی کارشناسی ارشد، گروه مهندسی آبیاری و آبادانی، پردیس کشاورزی و منابع طبیعی دانشگاه تهران، کرج، ایران.

۲. استادیار، گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه کردستان، سنندج، ایران.

۳. استاد، گروه مهندسی آبیاری و آبادانی، پردیس کشاورزی و منابع طبیعی دانشگاه تهران، کرج، ایران.

تاریخ پذیرش مقاله: ۱۴۰۰/۰۵/۱۵

تاریخ دریافت مقاله: ۱۴۰۰/۰۲/۲۴

چکیده

در این پژوهش عملکرد روش‌های ماشین‌برداری پشتیبان و شبکه عصبی تابع پایه شعاعی در پیش‌بینی کیفیت آب سیمینه‌رود مقایسه شده است. برای این منظور پارامترهای نسبت جذب سدیم و یون کلر به‌عنوان شاخص‌های کیفیت آب در مصارف کشاورزی در نظر گرفته شد. از داده‌های اندازه‌گیری شده یون سدیم، کلسیم، منیزیم، pH، EC و دبی جریان به‌عنوان ورودی مدل‌ها طی یک دوره آماری ۱۲ ساله (۱۳۹۳-۱۳۸۲) در مقیاس ماهانه استفاده شد. ارزیابی نتایج براساس معیارهای ضریب همبستگی، جذر میانگین مربعات خطا و میانگین خطای مطلق انجام شد. نتایج دوره صحت‌سنجی در چهار ایستگاه پل بوکان، داشبند بوکان، قزل‌گنبد و کاولان نشان داد که مدل ماشین بردار پشتیبان در مقایسه با شبکه عصبی تابع پایه شعاعی، دارای ضریب همبستگی بالاتر (SVM: ۰/۷۱ تا ۰/۹۴، RBF: ۰/۳ تا ۰/۵)، ریشه میانگین مربعات خطا کم‌تر (SVM: ۰/۰۲۸ تا ۰/۰۷۵، RBF: ۰/۰۶۷۲ تا ۰/۰۳۱۷)، خطای میانگین مطلق کم‌تر (SVM: ۰/۰۰۳ تا ۰/۰۳۳، RBF: ۰/۰۸۷ تا ۰/۰۱۹) و برای پارامتر یون کلر و با همان ترتیب مقادیر SVM: ۰/۶۳ تا ۰/۸۸، RBF: ۰/۲۱ تا ۰/۳۸، SVM: ۰/۰۱۳ تا ۰/۰۸۲ و RBF: ۰/۰۱۴۷ تا ۰/۰۲۵، SVM: ۰/۰۰۸۵ تا ۰/۰۴۶ و RBF: ۰/۰۶۵۳ تا ۰/۰۹۹۶ mg/l برای نسبت جذب سدیم است. لذا براساس نتایج مدل ماشین بردار پشتیبان نسبت به مدل شبکه عصبی تابع پایه شعاعی از دقت و عملکرد بهتری برای پیش‌بینی پارامترهای کیفیت آب رودخانه سیمینه‌رود برخوردار است.

کلیدواژه‌ها: پارامترهای کیفی، مدل‌سازی، منابع آب، هوش مصنوعی.

Comparison of the support vector machine and radial function neural network models in predicting of SiminehRood river water quality, Iran

Bahareh Hosseinpahani¹, Saman Nikmehr², Kumars Ebrahimi^{3*}

1. M.Sc. Student, Department of Irrigation & Reclamation Engineering, College of Agriculture & Natural Resources, University of Tehran, Karaj, Iran.

2. Assistant Professor, Department of water sciences and engineering, faculty of agriculture, University of Kurdistan, Sannandaj, Iran.

3. Professor, Department of Irrigation & Reclamation Engineering, College of Agriculture & Natural Resources, University of Tehran, Karaj, Iran.

Received: May 14, 2021

Accepted: August 06, 2021

Abstract

In this study, the performance of Support Vector Machine (SVM) and Radial Base Neural Network approach in predicting the water quality of SiminehRood River was examined. For this purpose, the Sodium Adsorption Ratio (SAR) and Chlorine ions were considered as indicators of water quality in agricultural use. Sodium, calcium, magnesium, pH, EC, and river flow rate were utilized as input monthly parameters throughout a 12-year period (2003-2014). The results evaluated based on correlation coefficient, root means square error and mean absolute error. The results of the validation period in 4 stations of Pol Bukan, Dashband Bukan, Ghezel Gonbad and Kaulan showed that the SVM model in comparison with the neural network of the radial base function, has higher correlation coefficient (SVM: 0.71 to 0.94, RBF: 0.3 to 0.5), the lowest root means square error (SVM: 0.028 to 0.075 mg/l, RBF: 0.0672 to 0.317 mg/l), a lower absolute mean error (SVM: 0.003 to 0.033 mg/l, RBF: 0.087 to 0.19 mg/l) for the chlorine ion parameter and in the same order SVM values: 0.63 to 0.88 and RBF: 0.21 to 0.38, SVM: 0.0013 to 0.0282 mg/l and RBF: 0.047 to 0.025 mg/l, SVM: 0.0085 to 0.046 mg/L and RBF: 0.0653 to 0.0996 mg/l for sodium absorption ratio. Therefore, the Support Vector Machine model has better accuracy and performance for predicting water quality parameters of SiminehRood River than the Radial Basis Function Network.

Keywords: Artificial intelligence, Modelling, Qualitative Parameters, Water resources.

مقدمه

کاهش منابع آب قابل استفاده و افزایش آلودگی آب رودخانه‌ها اثرات زیان‌باری بر بخش‌های مختلف مصرف به‌ویژه کشاورزی دارد. تأمین آب سالم یکی از بزرگ‌ترین چالش‌ها و اهداف مدیران و برنامه‌ریزان منابع آب است. تأمین آب شرب و کشاورزی از منابع آب سطحی در دسترس، تأثیر به‌سزایی در جلوگیری از هدررفت آب‌های سطحی و بهره‌مندی بیشتر از این منابع آب ارزشمند دارد. در راستای افزایش بهره‌وری از منابع آب سطحی ارزیابی وضعیت کیفی آن‌ها از اهمیت ویژه‌ای برخوردار است. ارزیابی کیفیت آب‌های سطحی، اثربخشی روش‌ها و رویکردها را در استفاده کمی از این منابع آبی تقویت نموده و سبب می‌شود تا زمینه‌های لازم برای تغییر و اصلاح تصمیم‌های بعدی فراهم آید. روش‌های هوش مصنوعی در زمینه‌های مختلف از جمله پیش‌بینی میزان آلودگی منابع آب نقش به‌سزایی دارد. افزایش مقادیر یون کلر سبب افزایش شوری آب می‌شود. هم‌چنین افزایش مقادیر پارامتر درصد جذب سدیم^۱ در آب آبیاری اثرات نامطلوبی هم‌چون تغییر خواص خاک، کاهش نفوذپذیری و عدم تهویه مناسب خاک را سبب می‌شود. بنابراین تخمین غلظت یون سدیم امری ضروری است (Kianian & Mohammadi, 2016). در همین راستا Ebrahimi (2018) به تخمین پارامتر EC رودخانه اهرچای با استفاده از شبکه عصبی مصنوعی و استنتاج فازی-عصبی تطبیقی پرداختند. نتایج مطالعه ایشان نشان داد که هر دو مدل با دقت بسیار خوبی هدایت الکتریکی را شبیه‌سازی می‌کند. Isazadeh et al. (2019) نیز پارامترهای کیفیت آبخوان دشت گیلان را با استفاده از آزمون گاما و مدل‌های ماشین بردار پشتیبان^۲ و شبکه عصبی مصنوعی^۳ بررسی کردند. نتایج حاصل از مطالعه اخیر حاکی از آن است که مدل ماشین بردار پشتیبان

عملکرد و دقت بیشتری را نسبت به مدل شبکه عصبی مصنوعی داشته و آزمون گاما توانسته است به‌صورت دقیق در تعیین ترکیب ورودی‌ها مؤثر مورد استفاده قرار گیرد. Aldhyani et al. (2020)، در مطالعه‌ای بیان داشتند که مدل‌سازی و پیش‌بینی کیفیت آب برای حفاظت از محیط‌زیست بسیار مهم است. لذا به بررسی استفاده از الگوریتم‌های پیشرفته هوش مصنوعی برای اندازه‌گیری کیفیت آب در آینده پرداختند. در این روش پیشنهادی، از الگوریتم‌های پیشرفته هوش مصنوعی، یعنی NARNET و LSTM برای پیش‌بینی WQI و برای طبقه‌بندی داده‌های WQI از الگوریتم‌های یادگیری ماشین مانند SVM، KNN و Naive Bayes استفاده کردند. نتایج این مطالعه نشان داد که عملکرد مدل NARNET براساس مقدار R به‌دست‌آمده کمی بهتر از مدل LSTM است. با این حال، الگوریتم SVM بالاترین دقت در پیش‌بینی WQC را در مقایسه با الگوریتم‌های KNN و Naive Bayes به‌دست آورده است. Salavati et al. (2016)، در مطالعه‌ای کاربرد تلفیقی مدل‌های برنامه‌ریزی پویا و ماشین بردار پشتیبان را در بهینه‌سازی بهره‌برداری از مخزن سد شاهچراغی در جهت حداقل‌سازی کمبود در طول دوره بهره‌برداری را بررسی کردند. نتایج این مطالعه حاکی از آن است که مدل ماشین بردار پشتیبان از عملکرد قابل‌قبولی در حل مسائل مربوط به بهره‌برداری از مخازن برخوردار است. از روش‌های هوشمند جهت شبیه‌سازی و پیش‌بینی پارامترهای کیفیت آب با توانایی زیاد به‌جای روش‌های تحلیلی می‌توان استفاده کرد. روش‌های سنتی پیش‌بینی کیفیت آب دارای محدودیت‌هایی هم‌چون محاسبات دشوار ریاضیات، رفتار یکسان با داده‌های قدیمی و جدید و دشواری پیش‌بینی در صورت هم‌پوشانی چند متغیره هستند (Xin et al., 2015). تجزیه و تحلیل کیفیت آب یک گام اساسی در مدیریت منابع آب است و برای کنترل

توانایی مدل ارائه‌شده در تعیین زمان پیش‌هشدار سیلاب در شرایط مختلف کاربری است. Ehteshami *et al.* (2016)، از دو شبکه عصبی پسانتشارخطا و تابع پایه شعاعی برای تخمین نیترات آب زیرزمینی بابل استفاده کردند. نتایج پژوهش ایشان نشان داد که هر دو شبکه نتایج قابل‌قبولی را در پیش‌بینی نیترات آب زیرزمینی داشته و می‌توان به‌عنوان ابزاری مناسب برای شبیه‌سازی پارامترهای ذکر شده استفاده کرد. Abbasian & Shahraki (2020)، با استفاده از شبکه‌های عصبی مصنوعی GMDH و RBF تقاضای ماهانه آب شهری زاهدان در سال ۱۳۹۶ را برآورد کردند. نتایج به‌دست‌آمده و مقایسه شاخص‌های MSE و MEA نشان داد که شبکه عصبی GMDH با سه لایه نهان که دارای یک نرون در لایه اول، سه نرون در لایه نهان دوم و سه نرون در لایه نهان سوم است، بهترین نتیجه را برای پیش‌بینی تقاضای کوتاه‌مدت آب داشتند. همچنین در لایه خروجی مدل‌های عصبی توابع غیرخطی عملکرد بهتری نسبت به توابع خطی از خود نشان داد. Nikpour & MahmodiBabelan (2019)، از مدل‌های شبکه عصبی مصنوعی، برنامه‌ریزی بیان ژن، موجک-عصبی و حداقل مربعات ماشین بردار پشتیبان به‌منظور روندیابی جریان روزانه رودخانه بالخلوچای و برای ارزیابی مدل‌ها از شاخص‌های آماری ریشه میانگین مربعات خطا، ضریب همبستگی، ضریب نش‌ساتکلیف و بایاس استفاده کردند. نتایج گویای عملکرد قابل‌قبول هر چهار مدل و برتری مدل حداقل مربعات ماشین بردار پشتیبان بود. اما در برآورد دبی بیشینه مدل موجک-عصبی خطای کم‌تری داشت. Shahinejad & Dehghani (2017)، رسوبات رودخانه کاکارضا را با استفاده از مدل ماشین بردار پشتیبان و برنامه‌ریزی بیان ژن تخمین زدند. نتایج پژوهش اخیر نشان داد که مدل ماشین بردار پشتیبان توانایی بالایی در برآورد مقادیر حداقل و حداکثر دارد.

هرگونه آلودگی که ممکن است بر اکوسیستم تأثیر منفی بگذارد و اطمینان از رعایت استانداردهای زیست‌محیطی، باید به آن توجه شود. این کار تلاشی برای توسعه یک مدل کارآمد با استفاده از ماشین بردار پشتیبانی برای پیش‌بینی کیفیت آب حوضه رودخانه لنگات^۴ توسط Abobakr Yahya *et al.* (2019)، از طریق تجزیه و تحلیل داده‌های شش پارامتر pH، جامدات معلق (SS)، اکسیژن محلول (DO)، نیتروژن آمونیاک (AN)، تقاضای اکسیژن شیمیایی (COD) و تقاضای اکسیژن بیوشیمیایی (BOD) مخازن دوگانه که در حوضه آبریز واقع شده، بوده است. نتایج این تحقیق نشان‌گر توانایی استفاده از SVM در پیش‌بینی خصوصیات کیفیت آب براساس این شش پارامتر است. Fathian & Hormozinezhad (2012)، با استفاده از سه مدل شبکه عصبی مصنوعی پرسپترون چندلایه، شبکه پیش‌رونده و شبکه تابع پایه شعاعی^۵، مقادیر سه روز آینده پارامترهای دبی، اشل، کل املاح موجود در آب و هدایت الکتریکی رودخانه کارون را پیش‌بینی کردند. نتایج پژوهش ایشان نشان داد که مدل شبکه پیش‌رونده دقت بیش‌تری نسبت به دو مدل دیگر در پیش‌بینی پارامترهای کمی و کیفیت آب را دارد. Santin (2015)، با استفاده از روش شبکه عصبی به پیش‌بینی پارامترهای بیولوژیکی پساب ناشی از فاضلاب پرداخت و با حداکثر مقدار مجاز در استفاده از پساب‌های فاضلاب برای آبیاری گیاهان مقایسه نمود. Arabgol & Sartaj (2012)، به پیش‌بینی غلظت نیترات در آبخوان دشت اراک با استفاده از شبکه ماشین بردار پشتیبان پرداختند و نتایج نسبتاً مناسبی را از غلظت نیترات به‌دست آوردند. Banihabib & Arabi (2008)، از مدل شبکه عصبی مصنوعی برای تعیین زمان پیش‌هشدار سیلاب در حوضه آبخیز گلابدره-دربند برای پنج سناریو مختلف تغییر کاربری استفاده کردند. نتایج مطالعه ایشان حاکی از

سیمینه‌رود است. این دو پارامتر از عوامل اصلی تأثیرگذار بر کیفیت منابع آب کشاورزی هستند و پیش‌بینی آن‌ها نقش مهمی در مدیریت منابع آب یک منطقه دارد.

مواد و روش‌ها

منطقه مورد مطالعه

حوضه آبریز سیمینه‌رود در استان آذربایجان غربی با مختصات جغرافیایی $35^{\circ} 45'$ تا $25^{\circ} 46'$ طول شرقی و $10^{\circ} 36'$ تا $56^{\circ} 37'$ عرض شمالی واقع شده است. رودخانه سیمینه‌رود از ارتفاعات ابراهیم جلال (۲۲۳۶ متر)، نیستان (۲۴۱۰ متر) و کپری (۲۴۱۷ متر) استان کردستان سرچشمه می‌گیرد. طول سیمینه‌رود ۲۰۰ کیلومتر، مساحت حوضه‌ی آن ۷۶۹ کیلومترمربع است. مشخصات و موقعیت ایستگاه‌های مورد مطالعه این رودخانه به ترتیب در جدول (۱) و شکل (۱) ارائه شده است.

به ارزیابی کاربرد آنتروپی شانون در تعیین کیفیت آب رودخانه‌های کارون و بابلرود پرداختند. نتایج این مطالعه نشان داد که آنتروپی شانون به خوبی وزن‌دهی پارامترهای کیفیت آب و تعیین شاخص کیفیت آب را انجام می‌دهد. Pirali Zefreh *et al.* (2020)، به شبیه‌سازی پارامتر کیفی کلروفیل و شفافیت آب در دوره زمانی ۳۲ ساله برای تالاب بین‌المللی چغاخور با استفاده از مدل ماشین بردار پشتیبان پرداختند. نتایج این مطالعه نشان داد که ماشین بردار پشتیبان قابلیت بالایی در پیش‌بینی کلروفیل و شفافیت آب دارد و با حداقل ورودی نیز قدرت پیش‌بینی لازم را داشته است.

در همین راستا و با توجه به اهمیت بالای بخش کشاورزی، پژوهش حاضر با هدف بررسی توانایی مدل‌های ماشین بردار پشتیبان و شبکه عصبی تابع پایه شعاعی در پیش‌بینی پارامترهای CI و SAR در رودخانه

Table 1. Characteristics of SiminehRood's Stations

Row	Stations	Longitude	Latitude	Height (meters)	Average annual (MCM)	Average flow (m ³ /s)
1	Pol Bukan	46°-11'-48"	36°-31'-04"	1328	283.8	9.00
2	Dashband Bukan	46°-10'-05"	36°-38'-44"	1311	452.2	14.34
3	Ghezel Gonbad	45°-57'-29"	36°-25'-32"	1372	162.4	5.15
4	Kaullan	45°-40'-44"	36°-23'-52"	1520	NA	NA

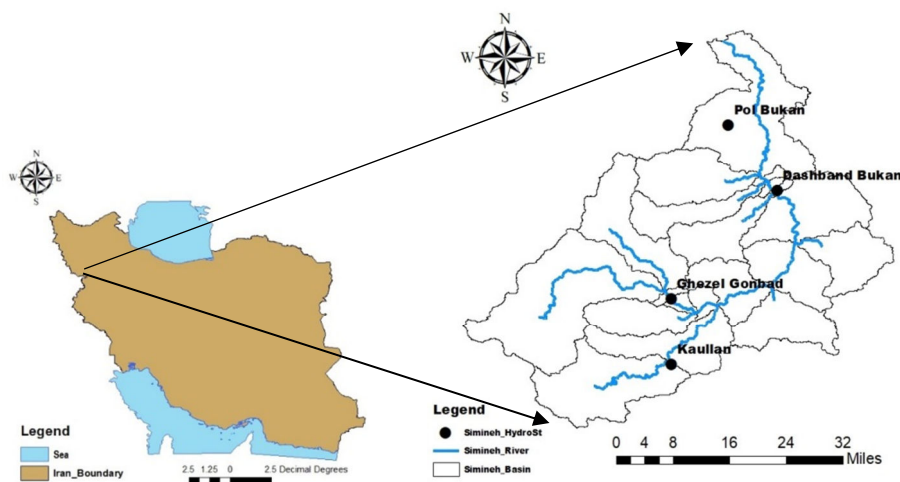


Figure 1. Location of SiminehRood stations

داده‌های مورد استفاده

در پژوهش حاضر، داده‌های رودخانه مورد مطالعه در مقیاس زمانی ماهانه طی سال‌های ۱۳۸۲ تا ۱۳۹۳ از شرکت سهامی آب منطقه‌ای کردستان تهیه و در حالت نرمال به‌عنوان ورودی برای مدل‌های SVM و RBF در دو دوره واسنجی و صحت‌سنجی مورد ارزیابی قرار گرفت. یون سدیم، کلسیم، منیزیم، pH، EC و دبی جریان رودخانه به‌عنوان ورودی هم‌چنین یون کلر و نسبت جذب سدیم به‌عنوان خروجی مدل‌های ماشین بردار پشتیبان و شبکه عصبی تابع پایه شعاعی معرفی شدند. مطابق معمول، ۷۰ درصد داده‌ها برای آموزش و ۳۰ درصد آن‌ها برای ارزیابی مدل‌ها استفاده شد. تقسیم بندی داده‌ها به‌صورت تصادفی بوده است. در جدول (۲)، ویژگی‌های آماری پارامترهای کمی و کیفی ایستگاه کاولان به‌عنوان نمونه در بازه زمانی مذکور ارائه شده است.

ماشین بردار پشتیبان (SVM)

نظریه ماشین بردار پشتیبان اولین بار توسط Vapnik (1995) ارائه شد. این تئوری بعدها توسط Vapnik (1998) به‌عنوان

ابزاری قدرتمند جهت طبقه‌بندی داده‌ها در علوم مختلف مورد استفاده قرار گرفت. ماشین بردار پشتیبان همانند دیگر روش‌های هوش مصنوعی بر مبنای الگوریتم داده‌کاوی عمل می‌کند. مهم‌ترین عملکرد مدل ماشین بردار پشتیبان طبقه‌بندی و خطی‌سازی یا رگرسیون داده‌ها است. در واقع SVM برای این‌که داده‌های غیرخطی را از هم تفکیک کند باید از تابع‌های انتقال مختلف استفاده کند

(www.analyticsvidhya.com/blog/2014/10/support-vector-machine-simplified). Vapnik (1998)، نسخه‌ای از SVM را پیشنهاد دادند که به جای طبقه‌بندی، عمل رگرسیون‌گیری را انجام می‌دهد و به SVR معروف است. در این مدل خطی‌سازی داده‌ها به‌وسیله الگوریتم خطی‌سازی SVR- ϵ و با معرفی تابع ϵ -sensitive انجام می‌پذیرد. در یک الگوریتم خطی‌سازی SVR- ϵ ، با فرض مقادیر ورودی و خروجی x_i و y_i ، هدف یافتن تابعی است که کم‌ترین انحراف ϵ را از y_i داشته باشد. اگر داده‌های آموزش به تعداد k سری به‌صورت $(x_1, y_1), \dots, (x_k, y_k)$ موردنظر باشند، تابع خطی‌سازی به‌صورت رابطه (۱) است:

$$f(x) = (w, x) + b, (w, x) \in R^N, b \in R \quad (1)$$

Table 2. Statistical characteristics of the used parameters of Kaullan station

Parameter	Unit	Number of records	Average	Minimum	Maximum	Standard deviation	Skewness
Calibration							
Q	m ³ /s	101	32.82	0.04	276.28	42.13	3.84
EC	ds/m	101	411.5	180	1070	135.72	2.01
pH	-	101	7.76	7.1	8.5	0.23	0.7
Ca	ppm	101	2.85	1.6	5.6	0.66	1.49
Mg	ppm	101	0.97	0.2	2.8	0.46	1.62
Na	ppm	101	0.58	0.19	2.1	0.36	2.32
Cl	ppm	101	0.53	0.1	1.7	0.26	2.54
SAR	ppm	101	0.39	0.16	1.33	0.19	2.303
Validation							
Q	m ³ /s	43	15.66	0.06	136.64	21.52	4.59
EC	ds/m	43	378.32	240	760	76.96	2.87
pH	-	43	8.02	7.4	8.4	0.2	0.7
Ca	ppm	43	2.52	2	4	0.34	2.086
Mg	ppm	43	0.85	0.3	2	0.27	1.74
Na	ppm	43	0.5	0.2	1.8	0.26	3.65
Cl	ppm	43	0.49	0.2	1.4	0.2	2.667
SAR	ppm	43	0.37	0.18	1.04	0.15	3.064

تست است. یکی دیگر از مسائل مشکل در طراحی تفکیک‌کننده برای SVM چند کلاسه است. ماشین بردار پشتیبان نسبت به سایر روش‌های دسته بندی داده‌ها در صورتی که انتخاب حل‌کننده معادله درجه دوم و انتخاب کرنل مناسب انجام شوند، از دقت بسیار خوب و بالایی برخوردار است.

شبکه عصبی تابع پایه شعاعی (RBF)

در پژوهش حاضر از شبکه عصبی تابع پایه شعاعی جهت پیش‌بینی پارامترهای کیفیت آب سیمین‌رود استفاده شد. تابع مبتنی بر شعاع در واقع تابعی است که از توابع شعاعی به‌عنوان تابع کرنل استفاده می‌کند و مقدار آن فقط به فاصله ورودی از یک نقطه مرکزی ارتباط دارد. RBFها نوع خاصی از شبکه‌های عصبی هستند که بر پایه تئوری‌های درونیابی شکل گرفته‌اند. برخلاف شبکه‌های MLP که دارای لایه‌های متوالی متعددی هستند، این شبکه از سه لایه شامل ورودی، مخفی و خروجی تشکیل شده است. در لایه مخفی از تابع انتقال گوسی و در لایه خروجی از تابع انتقال خطی استفاده می‌شود. در واقع نرون روش RBF یک تابع گوسی است. ورودی این تابع، فاصله اقلیدسی بین هر ورودی به نرون با بردار مشخص شده هم‌اندازه بردار ورودی است. معادله (۴) فرم کلی نرون‌های خروجی در شبکه RBF را نمایش می‌دهد.

$$G_j(x) = \sum_{i=1}^k w_{ji} \phi(\|x - \mu_i\|; \sigma_i) \quad (4)$$

که در آن $G_j(x)$ تابع وابسته به زمین واحد خروجی و ترکیب خطی از k تابع پایه شعاعی ϕ با مرکز μ_i و پهنای باند σ_i و هم‌چنین w_{ji} وابسته به زمین کلاس و i امین مرکز است. تابع $\phi(\|x - \mu_i\|)$ را تابع پایه شعاعی گویند و $\|$ به معنی فاصله اقلیدسی است. در شبکه RBF فاصله هر الگو با بردار مرکز هر نرون در لایه میانی به‌عنوان تابع محرک شعاعی محاسبه می‌شود (Yilmaz & Keynar, 2011).

در تعیین تابع $f(x)$ باید همواره مقادیر خروجی به‌گونه‌ای تعیین شوند که به‌طور هم‌زمان میزان کم‌ترین انحراف ε تعیین شده و تابع مقدار مناسب w را اختیار کند. با حل رابطه بهینه‌سازی (۲) دست‌یابی به چنین شرایطی میسر می‌شود.

$$\text{Minimize } \frac{1}{2} w^2 + c \sum_{i=1}^k (\xi_i) \quad (2)$$

$$\left\{ \begin{array}{l} y_i - (w, x_i) - b \leq \varepsilon + \xi_i \\ (w, x_i) + b - y_i \leq \varepsilon + \xi_i \\ \xi_i, \xi'_i \geq 0 \text{ for all } \dots i = 1, 2, \dots, k \end{array} \right.$$

در رابطه (۲)، C که مقدار خطا است در واقع جایگزین را برای تابع $f(x)$ مشخص کرده و مقدار آن تا مقدار مجاز انحراف ε افزایش می‌یابد. درنهایت با استفاده از توابع کرنل $k(x_i, y_i)$ و پارامترهای لاگرانژ λ_i و λ'_i رابطه تابع خطی‌سازی به‌صورت رابطه (۳) بازنویسی می‌شود.

$$f(x, b) = \sum_{i=1}^k (\lambda'_i - \lambda_i) k(x_i - x) + b \quad (3)$$

در روش ماشین بردار پشتیبان قبل از خطی‌سازی داده‌ها، برای این که بتوان داده‌های با پیچیدگی بالا را پیش‌بینی نمود داده‌ها به‌وسیله تابع کرنل به فضای با ابعاد خیلی بالاتر منتقل می‌شوند (Komasi et al., 2017). با استفاده از توابع کرنل مختلف شامل چندجمله‌ای، خطی، توابع شعاعی پایه، درجه دو و پرسپترون چندلایه می‌توان آموزش و ارزیابی را انجام داد. تابع کرنل استفاده‌شده در این پژوهش، تابع شعاعی پایه است که با استفاده از کرنل RBF فضای ویژگی از طریق یک تبدیل غیرخطی به‌دست می‌آید.

ماشین بردار پشتیبان برخلاف شبکه‌های عصبی در ماکزیم‌های محلی گیر نمی‌افتد. برای داده‌های با ابعاد بالا تقریباً خوب جواب می‌دهد. بزرگ‌ترین محدودیت روش ماشین بردار پشتیبان در انتخاب کرنل است، زمانی که کرنل ثابت باشد، دسته‌بندی SVM تنها دارای یک پارامتر قابل تغییر (پارامتر جریمه خطا) توسط کاربر است. انتخاب بهترین کرنل برای یک مسئله خاص از مسائل مهم و مطرح است. دومین محدودیت در سرعت و اندازه در آموزش و

نتایج و بحث

با استفاده از نرم‌افزار تاناگرا و متلب مدل‌ها به‌گونه‌ای تنظیم شدند که در هر مرحله داده‌های CI و SAR در این دوره زمانی به‌عنوان خروجی به مدل معرفی و با نتایج حاصل از مدل‌سازی قابل مقایسه باشد. شکل‌های (۲) و (۳)، نمودار تغییرات غلظت و هم‌چنین عملکرد مدل‌های SVM و RBF را برای یون کلر به‌ترتیب در دوره واسنجی و صحت‌سنجی برای ایستگاه کاولان به‌عنوان نمونه‌ای از ایستگاه‌ها نشان می‌دهد. همان‌طور که شکل‌های (۲-A) و (۳-A) نشان می‌دهد، انطباق بسیار خوبی بین مقادیر پیش‌بینی شده یون کلر با مقادیر مشاهداتی در مدل SVM وجود دارد. در صورتی که مطابق شکل‌های (۲-B) و (۳-B) در دوره واسنجی و صحت‌سنجی انطباق کمی بین داده‌های مشاهداتی و پیش‌بینی شده توسط مدل RBF وجود دارد. بهترین نتایج بهینه مرحله آموزش و آزمون مدل‌ها برای یون کلر در جدول (۳) ارائه شده که مقادیر خطا و ضریب کارایی مدل‌ها را نشان می‌دهد. از آنجایی که هرچه مقدار ضریب تبیین به یک نزدیک‌تر باشد و مقدار خطا کم‌تر باشد عملکرد مدل بهتر است، مدل SVM در تخمین مقادیر یون کلر عملکرد خوب و نزدیک به مقدر واقعی داشته است.

با توجه به شکل (۲)، بیش‌ترین مقدار اختلاف بین مقادیر مشاهداتی و پیش‌بینی شده در دوره واسنجی مدل ماشین بردار ۰/۱۰۴ میلی‌گرم در لیتر و در مدل تابع پایه شعاعی برابر ۰/۳۵ میلی‌گرم در لیتر است. هم‌چنین همان‌طور که در شکل (۳) نمایان است در دوره صحت‌سنجی این مقادیر به‌ترتیب برابر با ۰/۱۶ و ۰/۳۳ میلی‌گرم در لیتر است.

پس از ارزیابی غلظت یون کلر، شاخص نسبت جذب سدیم به‌عنوان خروجی به مدل معرفی شد. این پارامتر در کیفیت آب مصرفی برای کشاورزی مؤثر بوده و افزایش غیرمجاز آن سبب کاهش نفوذ آب به خاک و تبادل هوای خاک می‌شود.

انتخاب نوع تابع انتقال از نکات مهم در مدل RBF است که بر عهده فرد مدل‌ساز است. در حالتی که نوع RBF، گوسی باشد، عرض آن تابع دارای اهمیت است و باید به‌گونه‌ای انتخاب شود که از فاصله میان نقاط داده بزرگ‌تر و از بعد داده‌ها کوچک‌تر باشد. یکی از موارد مهم دیگر در طراحی شبکه‌های RBF، تعداد مراکز مورد استفاده در لایه میانی است. انتخاب تعداد این مراکز رابطه مستقیم با دقت و پیچیدگی شبکه دارد، بنابراین این مراکز باید به‌گونه‌ای انتخاب شود که توازن میان دقت مورد نظر و پیچیدگی برقرار شود (Morshedy & Memarian, 2015).

آماره‌های ارزیابی مدل‌ها

به‌منظور ارزیابی نتایج مدل‌ها از سه معیار ریشه میانگین مربعات خطاها^۶، ضریب همبستگی^۷ و میانگین خطای مطلق^۸ که نشان‌گر میزان بازده شبکه است، استفاده شد (Rezaei & Mir Mohammadi Meybodi, 2014).

$$R^2 = \frac{\sum_{i=1}^n (y_{io} - y_{ip})^2}{\sum_{i=1}^n (y_{io} - \bar{y})^2} \quad (5)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_{io} - y_{ip})^2}{n}} \quad (6)$$

$$MAE = \frac{\sum_{i=1}^n |y_{io} - y_{ip}|}{n} \quad (7)$$

در روابط (۵)، (۶) و (۷)، y_{io} داده‌های مشاهداتی، y_{ip} داده‌های محاسباتی، \bar{y} متوسط داده‌های مشاهداتی و n تعداد کل داده‌ها است. RMSE ریشه میانگین مربعات خطا را نشان می‌دهد. MAE مجموع اختلافات مقدار اندازه‌گیری شده پارامترها از مقدار واقعی آن تقسیم بر تعداد کل داده‌ها است که اگر مقدار اندازه‌گیری شده از مقدار واقعی بزرگ‌تر باشد، علامت منفی شده و تقریب اضافی است. کم‌ترین مقدار RMSE بالاترین صحت پیش‌بینی را نشان می‌دهد. هرچه مقدار RMSE و MAE به صفر و R^2 به یک نزدیک‌تر باشد، جواب بهتری برای مدل حاصل شده است.

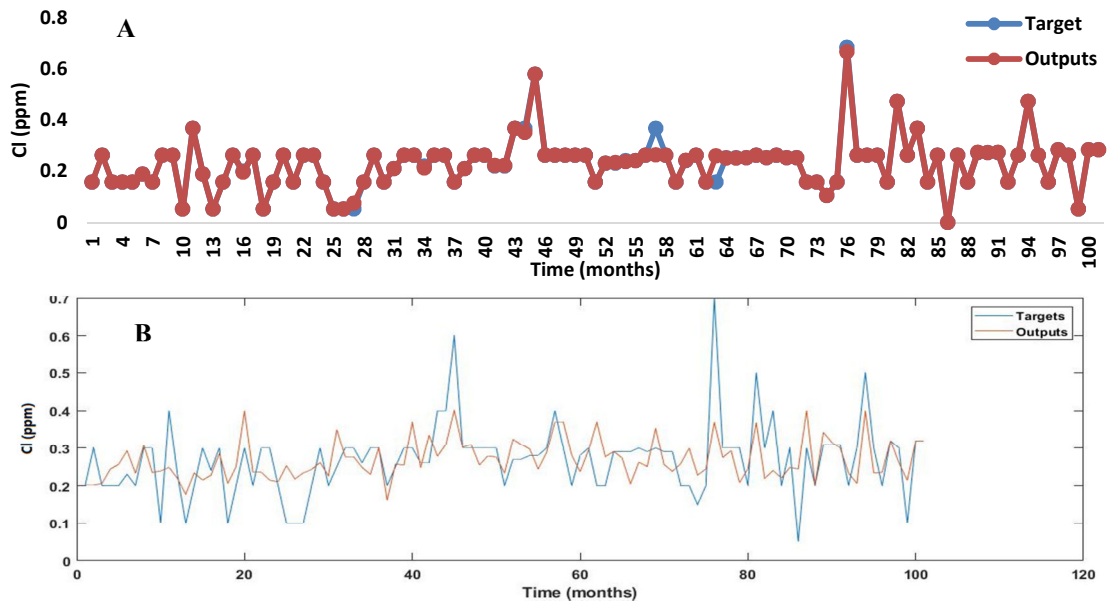


Figure 2. Performance of models in predicting Chlorine ion (mg /l), calibration period, (A) SVM model, (B) RBF model, Kaullan station

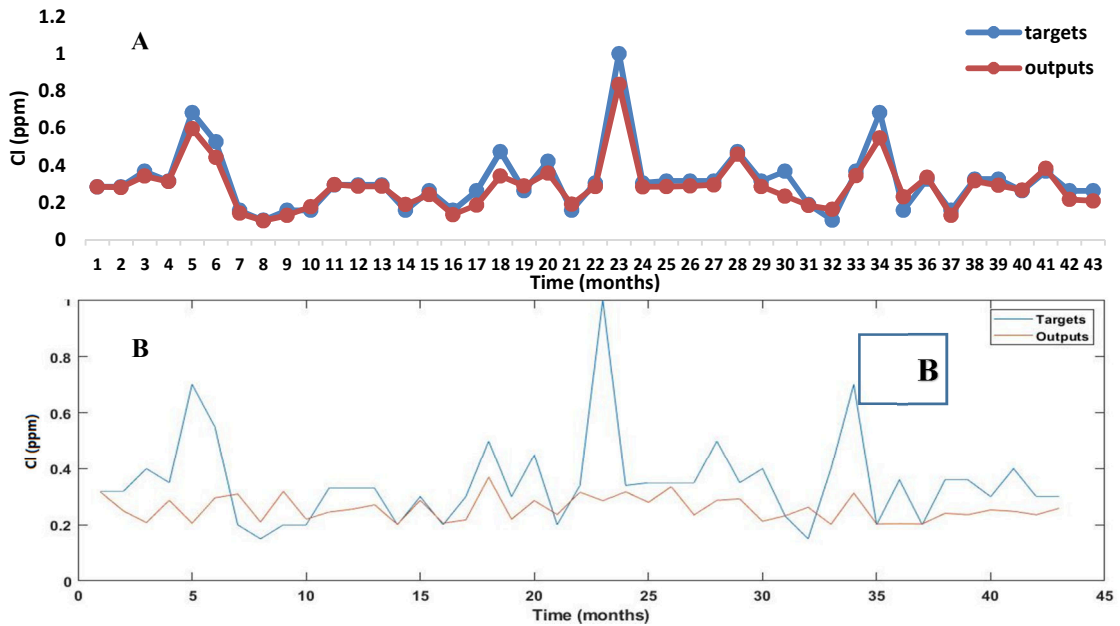


Figure 3. Performance of models in predicting Chlorine ion (mg/l), validation period, (A) SVM model, (B) RBF model, Kaullan station

Table 3. Evaluation indicators of models in calibration and validation periods for Chlorine ions

Station	Model	Calibration			Validation		
		R ²	RMSE (PPM)	ME (PPM)	R ²	RMSE (PPM)	ME (PPM)
Pol Bukan	SVM	0.99	0.00123	0.0001	0.81	0.039	0.012
	RBF	0.81	0.0073	0.051	0.45	0.14	0.087
Dashband Bukan	SVM	0.99	0.0014	0.00024	0.88	0.075	0.033
	RBF	0.91	0.006	0.044	0.5	0.48	0.11
Ghezel Gonbad	SVM	0.99	0.016	0.0017	0.71	0.028	0.003
	RBF	0.48	0.0076	0.0555	0.36	0.0672	0.19
Kaullan	SVM	0.97	0.014	0.000212	0.94	0.055	0.26
	RBF	0.31	0.0062	0.0584	0.3	0.317	0.11

مقادیر خطا و ضریب تبیین مدل‌ها حاکی از آن است که مدل ماشین بردار پشتیبان در تخمین مقادیر پارامترها عملکرد بهتری نسبت به مدل شبکه عصبی تابع پایه شعاعی در دو دوره واسنجی و صحت‌سنجی داشته است.

باتوجه به شکل (۴)، بیش‌ترین مقدار اختلاف بین مقادیر مشاهداتی و پیش‌بینی‌شده در دوره صحت‌سنجی مدل ماشین بردار ۰/۵۲ میلی‌گرم در لیتر و در مدل تابع پایه شعاعی برابر ۰/۴ میلی‌گرم در لیتر است.

شکل (۴)، عملکرد مدل‌ها را در پیش‌بینی پارامتر SAR در دوره صحت‌سنجی برای ایستگاه کاولان به‌عنوان نمونه‌ای از ایستگاه‌ها نشان می‌دهد. همان‌طور که در شکل (۴) نتایج هر دو مدل نسبت به مقادیر مشاهداتی در طی زمان نشان داده شده است، انطباق بسیار خوبی بین مقادیر پیش‌بینی‌شده نسبت جذب سدیم با مقادیر مشاهداتی در مدل SVM وجود دارد. جدول (۴)، بهترین نتایج بهینه مرحله آموزش و آزمون مدل‌ها برای نسبت جذب سدیم را نشان می‌دهد. هم‌چنین

Table 4. Evaluation indicators of models in calibration and validation periods for SAR

Station	Model	Calibration			Validation		
		R2	RMSE (PPM)	MAE (PPM)	R2	RMSE (PPM)	MAE (PPM)
Pol Bukan	SVM	0.99	0.0019	0.0002	0.88	0.05	0.0085
	RBF	0.84	0.0025	0.027	0.31	0.0147	0.0653
Dashband Bukan	SVM	0.99	0.00205	0.000039	0.71	0.082	0.021
	RBF	0.86	0.005	0.037	0.3	0.025	0.07
Ghezel Gonbad	SVM	0.99	0.0013	0.00003	0.86	0.045	0.014
	RBF	0.61	0.0045	0.0455	0.38	0.0148	0.0839
Kaullan	SVM	0.99	0.0018	0.000076	0.63	0.0013	0.046
	RBF	0.35	0.0045	0.043	0.21	0.0182	0.0996

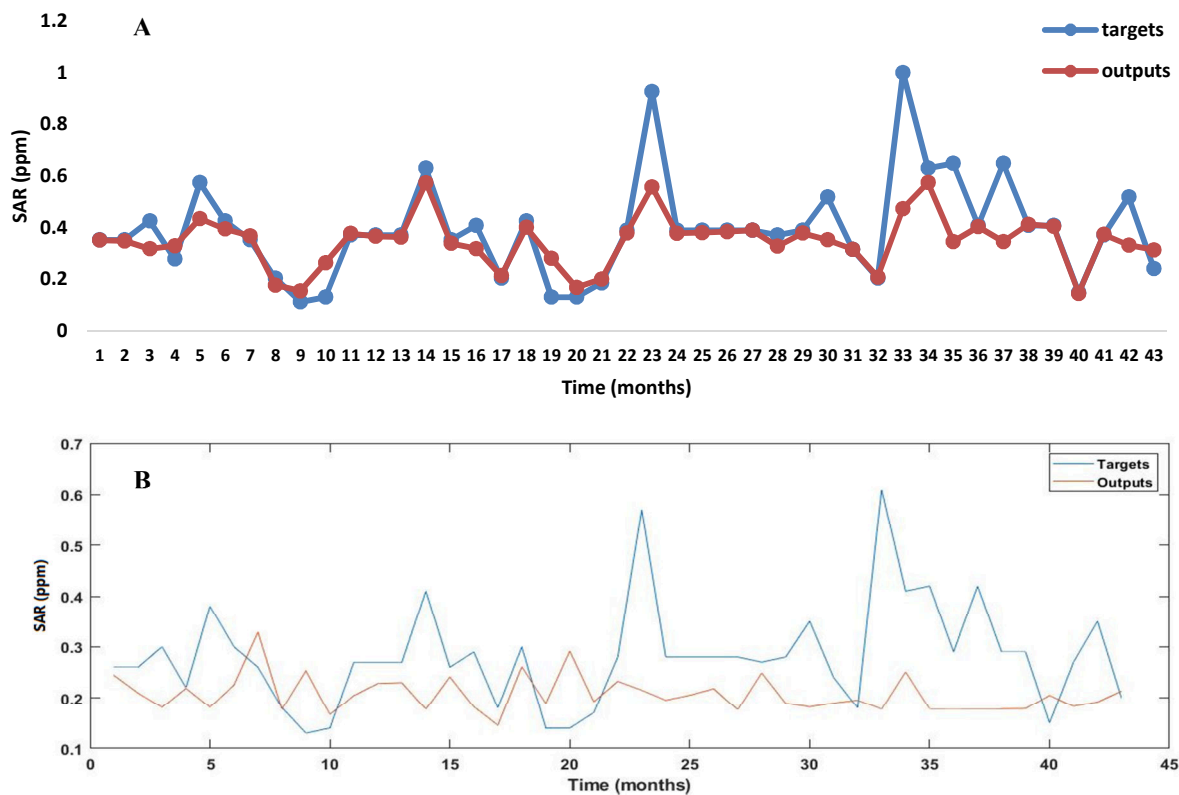


Figure 4. Performance of models in predicting SAR, validation period, (A) SVM model, (B) RBF model, Kaullan station

نتیجه‌گیری

در پژوهش حاضر به بررسی و مقایسه دو مدل ماشین بردار پشتیبان و شبکه عصبی تابع پایه شعاعی در پیش‌بینی پارامترهای کیفی رودخانه سیمینه‌رود پرداخته شد. برآورد پارامترهای کیفی آب با استفاده از روش‌های هوشمند کاهش هزینه‌ها و دسترسی سریع جهت مدیریت بهینه منابع آب را سبب می‌شود. برای پارامتر SAR، ضریب همبستگی در دوره واسنجی برای مدل SVM در همه ایستگاه‌های مورد مطالعه در مقایسه با مقادیر به‌دست‌آمده از مدل RBF مقدار قابل قبول ۰/۹۹ به‌دست آمد و در دوره ارزیابی این مقدار بین ۰/۶۳ تا ۰/۸۸ متغیر بود. همچنین مقدار ضریب همبستگی برای یون کلر مقدار ۰/۹۷ تا ۰/۹۹ در دوره واسنجی و ۰/۷۱ تا ۰/۹۴ در دوره صحت‌سنجی برای مدل SVM پیش‌بینی شد که در مقایسه با مقادیر به‌دست‌آمده از مدل RBF عملکرد بهتر SVM بارز است. مقدار جذر میانگین خطا و میانگین خطای مطلق در دوره واسنجی و ارزیابی برای مدل‌های SVM و RBF حاکی از دقت بهتر مدل ماشین بردار پشتیبان است. مدل ماشین بردار پشتیبان و شبکه عصبی مصنوعی هر دو از سرعت محاسبه، قابلیت اجرای ساده و دقت بالا برخوردار هستند. با توجه به این نکته که ماشین بردار پشتیبان نسبت به سایر شبکه‌های عصبی، تئوری قوی‌تری دارد و با ساختار برنامه‌ای درجه دوم خود به بهینه مطلق می‌رسد و همچنین نتایج به‌دست‌آمده، مدل SVM در پیش‌بینی پارامترهای کیفی عملکرد و دقت بهتری نسبت به مدل RBF داشته است. روش ارائه‌شده در این پژوهش را می‌توان به‌عنوان الگو برای پیش‌بینی دیگر پارامترهای کیفیت آب مورد استفاده قرار داد.

تشکر و قدردانی

انجام این پژوهش و تهیه مقاله‌های مربوطه با پشتیبانی

دانشگاه تهران، دانشگاه کردستان و سازمان آب منطقه‌ای

کردستان ممکن شد که بدینوسیله قدردانی می‌شود.

پی‌نوشت‌ها

1. Sodium Adsorption Ratio
2. Support Vector Machine
3. Artificial Neural Network
4. Langat River
5. Radial Basis Function
6. Root Mean Square Error
7. Correlation Coefficient
8. Mean Absolute Error

تعارض منافع

هیچ‌گونه تعارض منافع توسط نویسندگان وجود ندارد.

منابع

- Abbasian, M., & Shahraki, A. (2020). Modeling and comparison of GMDH and RBF artificial neural networks in predicting short-term drinking water demand in Zahedan. *Iranian Journal of Irrigation and Water Engineering*, 10 (3), 248-261. (In Persian).
- Abobakr Yahya, A.S., Ahmed, A.N., Binti Othman, F., Ibrahim, R. K., Afan, H. A., El-Shafie, A., Fai, C. M., Hossain, M. S., Ehteram, M., & Elshafie, A. (2019). Water quality prediction model based support vector machine model for ungauged river catchment under dual scenarios: *Water*, 11(6), p. 1231.
- Akhoni Pourhosseini, F., & Ebrahimi, K. (2019). Shannon's Entropy Evaluation on Determination of Surface Water Quality (Case Studies: Karun and Babolrood Rivers). *Iranian Journal of Water and Irrigation Management*, 9(2), 171-183. (In Persian).
- Aldhyani, T. H., Al-Yaari, M., Alkahtani, H., & Maashi, M. (2020). Water quality prediction using artificial intelligence algorithms: Applied Bionics and Biomechanics, v. 2020.
- Arabgol, R.F., & Sartaj, M. (2012). Evaluation of efficiency of support vector machines in estimating nitrate concentration in groundwater. *In: Sixth National Congress of Civil Engineering*, 26-27 April, Semnan University, Iran. (In Persian).
- Banihabib, M. E. & Arabi, A. (2008). Artificial Neural Network Model for Determining Flood Warning Time in Golabdereh-Darband Basin. *In: Third Iranian Water Resources Management Conference*, 14-16 Oct, Tabriz University, Iran. (In Persian).

- Ehteshami, M., Dolatabadi Farahani, N., & Tavassoli S. (2016). Simulation of nitrate contamination in groundwater using artificial neural networks. *Modeling Earth Systems and Environment*, 28, 10(2).
- Fathian, H., & Hormozinezhad, A. (2012). Prediction of quantitative and qualitative parameters of Karun river flow using artificial neural network. *Journal of Wetland*, Islamic Azad University of Ahvaz, (8)5, 43-29. (In Persian).
- Isazadeh, M., Biazar, S., Ashrafzadeh, A., & Khanjani, R. (2019). Estimation of Aquifer Qualitative Parameters in Guilans Plain Using Gamma Test and Support Vector Machine and Artificial Neural Network Models. *Environmental Science and Technology*, 21(2), 1-21. (In Persian).
- Kianian, A., Mobarghaee Dinan, N., & Hashemi, H. (2016). Zoning of soils by irrigation with sewage using interpolation method (IDW) (Case study: southern city of Ray). *Journal of Environmental Research*, 7 (14), 81-90. (In Persian).
- Komasi, M., Goodarzi, H., & Behnia, A. (2017). Investigation of spatial-temporal fluctuations of groundwater water table by support vector and kriging machine (IDW). *Journal of Soil and Water Conservation Science (Agriculture and Natural Resources)*, 24(4), 71-80. (In Persian).
- Mohammadi, P., & Ebrahimi, K. (2018). Estimation of electrical conductivity of Aharchai River using neural network models and adaptive neural-fuzzy inference. In: *National Hydraulic Conference of Iran*, 4-6 Sep, Shahre kord University, Iran. (In Persian).
- Morshedy, A., & Memarian, H. (2015). A New Method of Generalized Radial Basis Function Network to Interpolate Regional Variables in Geosciences. *Scientific Quarterly Journal, Geosciences*, 24 (96), 107-116.
- Nikpour, M., & MahmodiBabelan, S. (2019). Compare the performance intelligent routing models daily river flow (Case study: River Balkhlouchay, Ardabil). *Iranian Journal of Irrigation and Water Engineering*, 8 (32), 64-78. (In Persian).
- Pirali Zefreh E, A.R., Hedayati, A., Pourmanafi, S., Beyraghdar, O., & Ghorbani,R. (2020). Evaluation of the efficiency of support vector machine in predicting changes in water quality parameters (Case study: Choghakhor International Wetland). *Iranian Journal of Aquatic Ecology*, 10(1), 23-34. (In Persian).
- Rezaei, A., & Mirmohammadi Meybodi, S.A.M. (2014). *Statistics and Probability: used in agriculture*. Isfahan: Academic Center for Education, Culture and Research, Isfahan University of Technology Press. (In Persian).
- Salavati, A., Banihabib, M. E., & Soltani, J. (2016). Hybrid Model for Reservoir Operation Optimization. In: *Water sciences and Engineering Conference*, 8-9 June, Shahid Beheshti Conference center, Tehran, Iran. (In Persian).
- Santin, I. (2015). Effluent Predictions in Wastewater Treatment Plants for the Control Strategies Selection, *Journal of Bilbao*, 2, 1009-1016.
- Shahinejad, B., & Dehghani, R. (2017). Evaluation and Performance of Support Vector Machine Model in Estimation of Suspended Sediment. *Journal of Irrigation and Water Engineering*, 8(1), 30-42. (In Persian).
- Vapnik,V. (1995). *The Nature of statistical learning Theory*. New York: Springer Press.
- Vapnik,V. (1998). *Statistical learning Theory*. New York, NY, USA: John Wiley Press.
- www.analyticsvidhya.com/blog/2014/10/support-vector-machine-simplified (Last access 2021 May 14)
- Xin, X., Li, K., Finlayson, B., & Yin, W. (2015). Evaluation, prediction and protection of water quality in Danjiangkou Reservoir. *Water science and Engineering*, 8, 30-39.
- Yilmaz, I., & Keynar, O. (2011). Multiple regression, ANN (RBF, MLP) and ANFIS models for prediction of swell potential of clayey soils. *Expert Systems with Applications*, 38(5), 5958-5966.