شاپا چاپی: ۲۲۵۱-۷۴۸۰
شاپا الکترونیکی: ۲۲۵۰-۷۴۰۰

نشریه حفاظت منابع آب و خاک

آدرس تارنما:
https://wsrcj.srbiau.ac.ir

پست الکترونیک:
iauwsrcj@srbiau.ac.ir
iauwsrcj@gmail.com

سال یازدهم
شماره چهار
تابستان ۱۴۰۱

تاریخ دریافت:
۱۴۰۰/۰۷/۰۷

تاریخ پذیرش:
۱۴۰۰/۱۰/۰۸

صفحات: ۱۲۷-۱۱۵

مدل‌سازی رفتار سدهای بتنی با استفاده از روش‌های شبکه عصبی مصنوعی و رگرسیون لجستیک

فردین سعید[۱]، محسن ایران دوست[۲]* و نوید جلال کمالی[۳]

۱) دانشجوی دکتری، گروه مهندسی آب، واحد کرمان، دانشگاه آزاد اسلامی، کرمان، ایران.
۲) استادیار، گروه مهندسی آب، واحد کرمان، دانشگاه آزاد اسلامی، کرمان، ایران.
۳) استادیار، گروه مهندسی آب، واحد کرمان، دانشگاه آزاد اسلامی، کرمان، ایران.
* ایمیل نویسنده مسئول: mohsenirandost@yahoo.com

چکیده:

زمینه و هدف: اندازه‌گیری و رفتارسنجی سد موضوع جدیدی می‌باشد که می‌تواند به دلیل تغییر پارامترهای موجود برای ارائه مدلی باشد که رفتار پارامترهای منفرد را بر روی سد و همچنین روی یکدیگر بررسی کند و تغییرات ایجاد شده را آنالیز کرده و سیاست‌های لازم را ایجاد کند. هدف این پژوهش ایجاد یک روش ترکیبی از رگرسیون لجستکی با بهینه‌سازی الگوریتم بهینه‌سازی گروه ذرات با مقدار حقیقی جهت پیش‌بینی رفتار تجهیزات سد می‌باشد.

روش پژوهش: در این مطالعه از داده‌های ۳۶۵ روز، از ۱۳۹۷/۰۱/۳۱ تا ۱۳۹۸/۰۱/۳۱ که ۶۰۰ مجموعه داده تجهیزات سد شامل پارامترهای دمای آب، سطح آب، فشار دریچه، میزان رسوب‌گذاری، فشار منافذ، دمای هوا، حجم آب ورودی، مشخصات ویژه سد، شرایط بتن، سطح آب مخزن، تغییر مکان افقی و عمودی، اجزای اتصال انتقالی و شتاب زمین، قدرت، فشار، کشش و تنش بالا برای مدل‌سازی استفاده شدند. برای آموزش مدل‌سازی الگوریتم بهینه‌سازی گروه ذرات مقدار حقیقی-رگرسیون لجستیک و ۱۲۰ مجموعه داده جهت آزمایش استفاده گردید. برای ارزیابی عملکرد روش پیشنهادی، از چهار آماره شامل ضریب تعیین (R^2)، ریشه میانگین مربعات خطا (RMSE)، ضریب پراکندگی (SI) و میانگین خطاهای انحرافی (MBE) استفاده گردید.

یافته‌ها: نتایج نشان داد که مدل در پیش‌بینی فشار پیزومتریک در بدنه سد عملکرد قابل قبولی دارد. همچنین نتایج حاصل از مدل شبکه عصبی مصنوعی با کسب R^2= ۰/۹۳۰ و SI=۸/۵۸۷ همگرایی قابل قبولی را نشان می‌دهد. نتایج مربوط به داده‌های آموزش مدل نیز بیانگر این است که میانگین (μ) و انحراف معیار (σ) مدل ارائه شده، برای داده‌های آموزش به ترتیب برابر با ۱/۳۴۱ و ۱/۵۲۶ می‌باشد و برای داده‌های صحت‌سنجی این مقادیر برابر با ۱/۵۷۶ و ۲/۲۴۷ می‌باشد که این نشان از عملکرد خوب مدل پیشنهاد شده دارد. در معیار احتمال تجمعی، مدل پیشنهاد شده با P_{50}=۰/۹۴۰ و P_{90}=۱/۷۴۲ مبین این است که نتایج قابل قبول می‌باشد.

نتایج: نتایج بیانگر این است که بهینه‌سازی گروه ذرات مقدار حقیقی-رگرسیون لجستیک در عوض به حداقل رساندن ریسک تجربی که تعمیم عالی برای اندازه‌های نمونه کوچک را فراهم می‌سازد، اصل کاهش ریسک سازه‌ای را اجرا می‌کند. نسبت مقادیر پیزومتریک پیش‌بینی‌شده به مقادیر قرائت‌شده برای حدود ۷۲ درصد داده‌ها در این مدل در حدود یک بوده که بیانگر آموزش و قدرت پیش‌بینی مناسب این مدل می‌باشد. در نهایت بر اساس معیارهای ارزیابی مدل ترکیبی عملکرد بهتری نسبت به روش‌های بیان شده دارد.

کلیدواژه‌ها: سد بتنی، تحلیل رگرسیون، الگوریتم بهینه‌سازی گروه ذرات، سد ستارخان

10.30495/WSRCJ.2022.19230

مقدمه

در یک حوضه آبریز سیستم‌های مختلف طبیعی و انسانی در اندرکنش با یکدیگر هستند که این امر موجب شده تا برنامه‌های مدیریت حوضه‌های آبریز اهمیت روز افزونی پیدا کنند. در مدیریت منابع آب در سطح حوضه آبریز تأثیر اندرکنش اجزای درون هر سیستم و اندرکنش سیستم‌های مختلف با یکدیگر باید در نظر گرفته شود. سدها دارای تعاملات قوی با عوامل محیطی، هیدرولیکی و ژئومکانیک، مانند درجه حرارت هوا و آب، سطح آب، فشار منافذ و شکل‌پذیری هستند که هر کدام بر ساختار و رفتار سد تأثیر می‌گذارد (Meier and Kramer, 2019; Neshat *et al*., 2016). حوضه‌های آبریز شامل عواملی همچون بارندگی، جریان آب سطحی، شکل حوضه، خصوصیات فیزیکی حوضه، دما، تبخیر، شرایط آب‌های زیرزمینی و ... با اقلیم زمان و شرایط آب و هوایی متغیر می‌باشد که هر کدام از این عوامل می‌تواند نقش تعیین کننده‌ای را در حالت کلی حوضه ایفا کند، از طرف دیگر به دلیل محدودیت منابع آب (بخصوص در ایران) لازم است که رفتار سدها در شرایط مختلف زمانی جهت مدیریت بهتر و استفاده بهینه و پارامترهای مؤثر بر آن بررسی شود. این امر نیازمند بررسی تک‌تک پارامترها و ارتباط متقابل آنها به‌صورت دینامیکی می‌باشد (Han *et al*., 2018). از سال ۱۹۳۰ تاکنون مدل‌های خطی و غیرخطی فراوانی جهت شبیه‌سازی و پیش‌بینی فرآیندها و متغیرهای هیدرولوژیکی پیشنهاد گردیده (Li *et al*., 2020) و این مدل‌ها در طول زمان پیشرفت کرده و به‌طور مداوم با بالا رفتن دانش بشری و معرفی ابزار جدید کیفیت آنها ارتقا یافته است. برای توصیف و پیش‌بینی رفتار ساختاری سدها، طی دهه‌های گذشته تعدادی از مدل‌های ریاضی قطعی، آماری و ترکیبی توسعه یافته است (Li and Wang, 2019; Reis *et al*, 2018; Samigulina and Massimkanova, 2019; Shakarami *et al*., 2018). در نتیجه، اندازه‌گیری و رفتار سد موضوع جدیدی می‌باشد که می‌تواند به دلیل تغییر پارامترهای موجود برای ارائه مدلی باشد که رفتار پارامترهای منفرد را بر روی سد و همچنین روی یکدیگر بررسی کند و تغییرات ایجاد شده را آنالیز کرده و سیاست‌های لازم را ایجاد کند. بدین منظور، چندین مدل ساخته شد که شامل مدل‌های کرک‌دار، شبکه‌های عصبی، مونت‌کارلو و ... می‌باشد. تجهیزات موجود در سد یک بسته نرم‌افزاری برای کنترل رفتار سد می‌باشد و به کاربر اجازه شبیه‌سازی را تحت روش سیستم طراحی دینامیکی را می‌دهد (Gitoee *et al*., 2018; Li *et al*., 2020). با استفاده از رگرسیون‌های معمولی در تجهیزات سد مدلی که به دست می‌آید که در موقعیت‌های گوناگون برای مصرف‌کنندگان حوضه است. از آنجایی که رفتار سد همانند رگرسیون به‌صورت قطعی نیست و تابع شرایط حاکم بر سد است نیازمند بهبود این تجهیزات با مدل‌های غیرقطعی مانند هوش مصنوعی و الگوریتم‌های فرا ابتکاری هستیم. مدل برنامه‌ریزی بهینه‌سازی نیز مدلی است که در مقایسه با روش مدل‌سازی سنتی هیچ نوع فرم تابع فرضی را از قبل در نظر نمی‌گیرد، اما به‌عنوان نمونه در روش رگرسیون، ساختار و سازه مدل از پیش تعیین شده است (مثلاً رگرسیون خطی، درجه دوم و...) و در مرحله بعدی پارامترهای مدل تعیین می‌شوند. هنگامی که برنامه‌ریزی بهینه‌سازی شده گروه ذرات معادله یا فرمولی بین متغیرهای ورودی و خروجی به دست می‌آورد، مزیت اصلی این روش توانایی انتخاب اتوماتیک متغیرهای ورودی در مدل مفید واقع شده و همچنین نادیده‌گرفتن متغیرهایی که مفید واقع نشده، می‌باشد (Schratz *et al*., 2018)؛ بنابراین، اغلب گفته می‌شود که اصلاح بهینه‌سازی گروه ذرات می‌تواند ابعاد متغیرهای ورودی را کاهش دهد و تعیین تغییرات در محیط‌های مختلف برنامه‌نویسی برای روابطی که ارائه می‌شود، امکان‌پذیر سازد. به عبارت دیگر، با تغییر نمودار هر پارامتر که بر سیستم تأثیرگذار می‌باشد، می‌توان رفتار سایر پارامترها را در نظر گرفت. علاوه بر این، رگرسیون لجستیک به دلیل تئوری آموزش آماری و یک اصل کاهنده ریسک سازه‌ای که با موفقیت برای

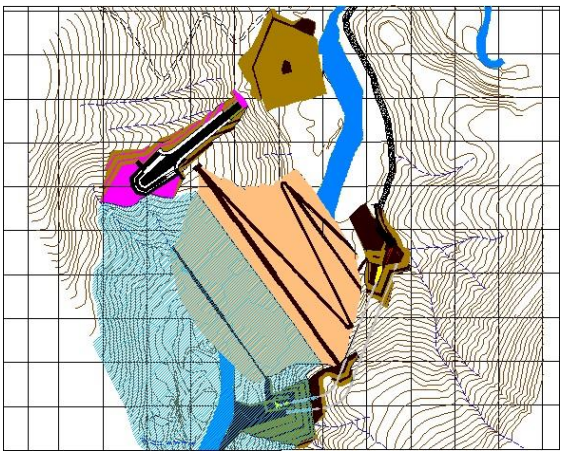
مدل‌سازی سیستم غیرخطی استفاده شد، یک ماشین آموزشی مؤثری باشد (Schratz *et al*., 2018). در مقایسه با شبکه‌های عصبی مصنوعی، رگرسیون لجستیک عملکرد قابل‌اعتمادتر و بهتری را در همان شرایط آموزشی ارائه می‌دهد (Kalita *et al*., 2020). با این حال ویژگی‌های بهتری دارا می‌باشد، اما رگرسیون لجستیک در تحقیقات آکادمیک و کاربردهای صنعتی به دلیل تعریف مناسب پارامترهای مختلف توسط کاربر محدود می‌باشد. پارامترهای رگرسیون لجستیک باید با دقت تنظیم شوند تا این مدل به‌طور مؤثرتری ساخته شود (Colkesen *et al*., 2016; Ma and Xia, 2017; Tsirikoglou *et al*., 2017). پارامترهای رگرسیون لجستیک به‌صورت نامناسب انتخاب شده که به دنبال آن یا تعداد بیش از حد متغیر مستقل در رگرسیون قرار داده شده یا تعداد کمتر از آن، در پی خواهد داشت و ممکن است تنظیمات مختلف پارامترها نیز باعث تفاوت قابل‌توجهی در عملکرد شوند (Siniscalchi and Salerno, 2016). بنابراین، انتخاب پارامترهای بهینه گامی مهم در طراحی رگرسیون لجستیک می‌باشد. با این وجود، هیچ نوع راهنمایی کلی برای کمک به انتخاب این پارامترها در دسترس نمی‌باشد (Acerbi and Ji, 2017; Oneto *et al*., 2016). امامی و همکاران (۱۳۹۹) با استفاده از الگوریتم فرا ابتکاری انتخابات (EA) و با لحاظ داده‌های تراز آب و دمای بتن طی سال‌های ۱۳۹۲-۱۳۷۹ به‌عنوان پارامترهای ورودی و مقدار تغییر مکان افقی و قائم ترک‌ها به‌عنوان پارامترهای خروجی، نحوه‌ی تغییرات ترک‌های سد بتنی قوسی زاینده‌رود را مورد ارزیابی قرار دادند و نتایج را با روش الگوریتم ژنتیک (GA) و شبکه‌ی عصبی مصنوعی (ANN) مقایسه کردند. برای ارزیابی عملکرد روش پیشنهادی، از سه آماره شامل ضریب تبیین(R^2) ، جذر میانگین مربعات خطا (RMSE) و معیار نش-ساتکلیف (NSE) استفاده گردید. نتایج آنها نشان داد که الگوریتم EA با کسب مقادیر R2=0.96، RMSE=0.022 و NSE=0.74 در مقایسه با دو روش الگوریتم GA و شبکه‌ی عصبی مصنوعی (ANN) از کارایی بالاتری برخوردار است. آنها همچنین بیان داشتند که برای ارزیابی تغییرات ترک‌های سدهای بتنی و پیش‌بینی روند تغییرات آن‌ها در آینده، الگوریتم‌های فرا ابتکاری روشی بسیار دقیق و قدرتمند محسوب شده و به‌وسیله این روش‌ها می‌توان دید بسیار خوبی بر وضعیت آسیب دیدگی سدهای بتنی پیدا نمود. Gu و همکاران (۲۰۲۰) یک رویکرد تحلیلی برای تغییر رفتار سدهای بتنی با استفاده از مدل‌های داده پانلی ارائه کردند. این رویکرد برای تشخیص اینکه آیا رفتار سدهای بتنی در طول عمر خود و خدمات آن دچار تغییر خواهند شد یا خیر، استفاده گردید. نتایج آنها بیانگر این است که روش بکار گرفته دارای اعتبار زمانی و مکانی در خصوص رفتار سدهای بتنی را دارد. Mata و همکاران (۲۰۲۱) اقدام به اعتبارسنجی مدل‌های یادگیری ماشین جهت مدل‌سازی رفتار سد سالاموند در پرتغال پرداختند. به این منظور از مدل‌های شبکه عصبی مصنوعی (ANN)، ماشین بردار پشتیبان (SVM)، جنگل تصادفی (RF) و درخت‌های رگرسیون تقویت شده (BRT) استفاده کردند. آنها بیان داشتند که قبل از استفاده از روش‌های ماشین یادگیرنده در ارزیابی و پیش‌بینی رفتار سدها ابتدا آنها بایستی از نظر کاربرد اعتبارسنجی شوند. روش‌های ماشین یادگیرنده ابزاری مناسب برای توسعه مدل‌های پایگاه داده هستند که عمدتاً زمانی بکار می‌روند که مقدار اطلاعات کافی موجود باشد و اعتبارسنجی آنها برای تحلیل و تفسیر رفتار سدها قابل اعتماد است. هدف این پژوهش ایجاد یک روش ترکیبی از رگرسیون لجستیک با بهینه‌سازی الگوریتم بهینه‌سازی گروه ذرات با مقدار حقیقی جهت پیش‌بینی رفتار تجهیزات سد می‌باشد. همچنین مدل مرسوم رگرسیون لجستیک و مدل شبکه عصبی نیز در رفتارسنجی سدهای بتنی مورد مقایسه قرار گرفتند.

مواد و روش‌ها

منطقه مورد مطالعه

سد مخزنی ستارخان واقع در استان آذربایجان‌شرقی در ۱۵ کیلومتری غرب شهرستان اهر و بر روی رودخانه ستارخان احداث شده است. رودخانه ستارخان از

ارتفاعات واقع در شمال‌غرب کشور در مرز ایران و ترکیه سرچشمه می‌گیرد و پس از عبور از شهر اهر و طی مسافتی در حدود ۵۰ کیلومتر به دریاچه ارومیه می‌ریزد. هدف از احداث سد ستارخان تأمین آب مورد نیاز شرب و صنعت شهر اهر و همچنین تأمین نیاز آب کشاورزی اراضی واقع در پایین دست سد ستارخان می‌باشد. سد ستارخان از نوع خاکی و بتنی- سنگریزه‌ای می‌باشد. تراز تاج سد در ارتفاع ۱۵۸۸ متر از سطح آزاد دریا و طول تاج حدود ۳۵۰ متر می‌باشد. مقطع خاکریز بدنه سد شامل هسته آب‌بند رسی، لایه‌های فیلتر و زهکش، پوسته‌های سنگریزه‌ای بالادست و پایین‌دست فرازبند که در داخل بدنه سد قرار می‌گیرد، می‌باشد. هسته آب‌بند رسی در مرکز مقطع سد بالاتر از تاج ۱۵۸۶ متر و با ارتفاع حداکثر ۱۱۶ متر از سنگ بستر قرار دارد که نسبت به محور سد متقارن می‌باشد. شیب شیروانی‌های بالادست و پایین‌دست هسته آب‌بند ۱:۳ (افقی به قائم) است. لایه فیلتر پایین دست به ضخامت افقی ۳ متر هسته آب‌بند رسی را از لایه زهکش جدا می‌سازد. لایه‌های زهکش بالادست و پایین-دست سد نیز با ضخامت افقی ۳ متر، آب خرابی تجهیزات سد شده از هسته آب‌بند را به خارج از بدنه سد تخلیه می‌نمایند. پوسته‌های بالادست و پائین‌دست به‌ترتیب با شیب‌های شیروانی ۲:۱/۵ (افقی به قائم) و شیب متوسط ۱:۱/۹ (افقی به قائم) از نوع سنگریزه‌ای تعیین شده‌اند.



شکل ۱. موقعیت سد ستارخان

شیب بالادست پوسته سد با لایه محافظ ریپ‌راپ و فیلتر زیر آن، در مقابل عمل امواج مخزن سد حفاظت می‌شوند، فرازبند سد با مقطعی مشابه مقطع بدنه اصلی سد، شامل هسته آب‌بند رسی در مرکز و پوسته‌های سنگریزه‌ای، در بالادست بدنه سد اصلی احداث می‌گردد. (شکل ۱) موقعیت سد ستارخان را نشان می‌دهد

بهینه‌سازی گروه ذرات

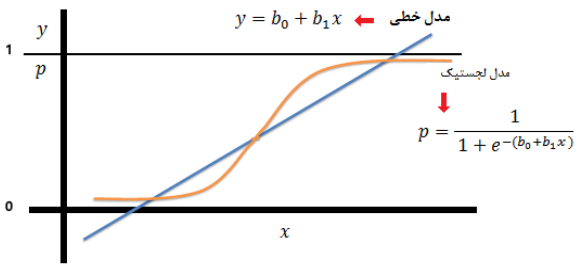
اساس حرکت ذرات، شبیه‌سازی رفتار جرمی می‌باشد که برای نشان‌دادن حرکت گروهی از پرندگان و ماهی‌ها استفاده می‌شود که یکی از رایج‌ترین تکنیک‌های بهینه‌سازی متاهیروستیک[1] (فرا ابتکاری) است. حرکت گروهی ذرات از فاکتور جمعیت، از جمله راه‌حل‌های بالقوه برای مسئله مورد مطالعه، برای کاوش در فضای جستجو استفاده می‌کند. در حرکت گروهی ذرات، هر یک از اعضای جمعیت دارای یک سرعت اتخاذی (جابه‌جایی) می‌باشد که با استفاده آن در فضای جستجو حرکت می‌کند. علاوه بر این، هر یک از آنها به دلیل دارا بودن یک حافظه، بهترین موقعیت در فضا را نیز دارا هستند. آنها تاریخچه جستجوی خود را به‌خاطر می‌آورند، در نتیجه هر یک از اعضا در دو جهت حرکت می‌کنند: به بهترین وضعیتی که تلاقی شده‌اند و به بهترین وضعیت بهترین عضوی که در همسایگی خود ملاقات کرده‌اند (Oneto *et al.*, 2016). یکی از جنبه‌های اصلی و جذابیت روش بهینه‌سازی گروه ذرات سادگی آن می‌باشد. به‌طوریکه فقط از دو معادله سرعت و مکان تشکیل شده است که در آن مختصات هر ذره پاسخ احتمالی مربوط به دو بردار را نشان می‌دهد. ظاهراً، $f: \mathbb{R}n \rightarrow \mathbb{R}$ تابع هزینه‌ای می‌باشد که باید به حداقل برسد. این تابع یک راه‌حل کاندیداتوری را به‌عنوان یک بردار از اعداد حقیقی به‌عنوان متغیر مستقل می‌گیرد و یک عدد حقیقی به‌عنوان خروجی تولید می‌کند که مقدار تابع هدف راه‌حل کاندیداتوری داده شده را نشان می‌دهد. شیب معین نمی‌باشد. هدف یافتن راه‌حل a برای f

[1] Metaheuristic

$f(a) \leq f(b)$ برای همه اهداف b در فضای جستجو می‌باشد که به عبارتی a حداقل جهانی می‌باشد. از طرفی S تعداد ذرات موجود در گروه می‌باشد، هرکدام دارای موقعیتی برابر با $xi \in \mathbb{R}n$ در فضای جستجو و سرعتی برابر با $vi \in \mathbb{R}n$ هستند. همین‌طور pi شناخته‌شده‌ترین موقعیت ذره i می‌باشد و g شناخته‌شده‌ترین موقعیت کل گروه می‌باشد (Acerbi and Ji, 2017).

رگرسیون لجستیک

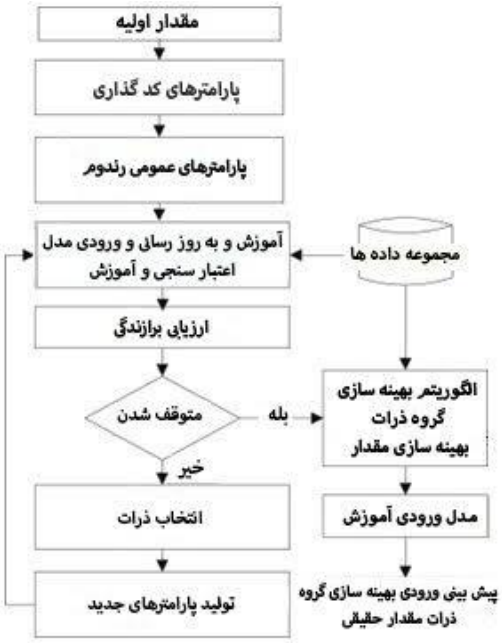
اغلب از ضریب همبستگی برای بیان شدت رابطه خطی بین دو متغیر کمی استفاده می‌شود. علاوه بر این، از مدل رگرسیون برای نشان دادن مدل رابطه بین این دو استفاده می‌کنیم. در این بین الگویی برای پیش‌بینی متغیر وابسته (Y) به دلیل متغیر مستقل(X) ایجاد می‌شود. با این وجود، در مدل ایجاد شده، هر دو متغیر مستقل و وابسته کوچک می‌باشند. همچنین، شرایط تداوم این مقادیر در روش رگرسیون نهفته می‌باشد. امکان دارد خواستار رابطه بین یک متغیر مستقل (با مقادیر پیوسته) و یک متغیر وابسته را با مقادیر کیفی بسنجیم (Hussain *et al.*, 2020). در این حالت، روش رگرسیون خطی نرمال کار نمی‌کند و باید از رگرسیون لجستیک استفاده شود. برای تعیین مدل رابطه بین متغیر وابسته و مستقل به‌جای رابطه خطی، به تابعی نیاز داریم که از ۰ تا ۱ متغیر باشد. روش رگرسیون لجستیک از تابعی به نام تابع لجستیک استفاده می‌کند. به همین خاطر، به این روش رگرسیون، رگرسیون لجستیک گفته می‌شود. در ادامه، این تابع تعریف شده و نمودار مربوط به آن در مورد پارامترهای موجود در (شکل ۲) قابل مشاهده می‌باشد.



شکل ۲. رگرسیون لجستیک

بهینه‌سازی مبتنی بر بهینه‌سازی گروه ذرات مقدار حقیقی مدل رگرسیون لجستیک

عملکرد تعمیم رگرسیون لجستیک با دقت برآورد و عملکرد بستگی به پارامتر ورود به سیستم دارا می‌باشد که به نحو احسن تنظیم شده است (Chambers, 2019; Feng *et al.*, 2019; Nan *et al.*, 2019). با این وجود، هیچ دستورالعمل کلی برای انتخاب این پارامترها در دسترس نمی‌باشد. در حال حاضر اغلب محققان یک روش استاندارد (آزمون و خطا) را با استفاده از الگوریتم شبکه دنبال می‌کنند، در وهله اول چندین مدل رگرسیون لجستیک را به دلیل مجموعه‌ای از پارامترهای مختلف ایجاد می‌کنند، در مرحله بعد، آنها را به دلیل بدست آوردن پارامترهای بهینه بر روی یک مجموعه اعتبارسنجی آزمایش می‌کنند (Belmokr *et al.*, 2019). بنابراین، در این مطالعه یک الگوریتم بهینه‌سازی‌شده گروه ذرات مقدار حقیقی را برای جستجوی پارامترهای بهینه رگرسیون لجستیک به‌منظور بهبود عملکرد پیش‌بینی، اتخاذ گردید. در مدل پیشنهادی الگوریتم بهینه‌سازی‌شده گروه ذرات مقدار حقیقی-رگرسیون لجستیک، مقادیر پارامترهای



شکل ۳. مدل الگوریتم بهینه‌سازی گروه ذرات مقدار حقیقی - رگرسیون لجستیک

رگرسیون لجستیک مستقیماً با داده‌های ارزش واقعی روی کروموزوم رمزگذاری می‌شوند. مقادیر پارامترهای رگرسیون لجستیک از طریق فرآیند تکامل الگوریتم بهینه‌سازی‌شده گروه ذرات مقدار حقیقی به‌صورت دینامیکی بهینه می‌شوند و از پارامترهای به‌دست‌آمده برای ساختن یک مدل رگرسیون لجستیک بهینه شده برای بهبود پیش‌بینی استفاده گردید. چارچوب بهینه‌سازی پارامترهای رگرسیون لجستیک با الگوریتم بهینه‌سازی گروه ذرات با مقدار حقیقی در (شکل ۳) نشان داده شده است.

گام ۱ (کدگذاری ذرات)

پارامترهای رگرسیون لجستیک برای تولید ذرات به‌صورت تصادفی و مستقیم کدگذاری می‌شوند (Chambers, 2019; Raj & Ananthi, 2019). دامنه به‌صورت‌های (۱۱۰۰)، (۰/۰۰۰۱، ۰/۰۱) و (۰، ۱) تعریف می‌شود. جمعیت حاوی ۱۰۰ ذره می‌باشد.

گام ۲ (تعریف برازندگی)

محاسبه برازندگی مجموعه داده‌های آموزشی آسان و تمایل به مناسب بودن می‌باشد. این مشکل با استفاده از یک روش اعتبارسنجی متقابل قابل‌حل می‌باشد. در این زمینه، برای غلبه بر پدیده بیش از حد مناسب، از یک روش اعتبار پنج‌گانه استفاده می‌شود (Rong *et al.*, 2019). تابع رگرسیون ساخته شده با مجموعه‌ای از پارامترهای داده شده می‌باشد و از چهار زیرمجموعه به‌عنوان یک مجموعه آموزش استفاده می‌کند. عملکرد مجموعه پارامترها با خطای ریشه میانگین مربع در زیرمجموعه آخر اندازه‌گیری می‌شود. روش فوق به‌صورت پنج بار تکرار می‌شود به‌طوریکه از هر زیرمجموعه یک‌بار برای اعتبارسنجی استفاده می‌شود. میانگین بهینه‌سازی گروه ذرات مقدار حقیقی در پنج امتحان آزمایشی خطای تعمیم مورد انتظار برای مجموعه‌های آموزشی می‌باشد.

گام ۳ (عملگرهای الگوریتم بهینه‌سازی‌شده گروه ذرات)

در عملگرها، از همگرایی استاندارد برای انتخاب ذرات عالی برای تولیدمثل استفاده می‌شود. یک همگرایی تک نقطه‌ای به‌طور تصادفی برای تبادل ذرات بین دو فاصله اتخاذ می‌شود: احتمال ایجاد ذره جدید برای هر جفت ۰/۵ می‌باشد. ضریب یک عمل همگرایی را دنبال می‌کند که تعیین می‌کند آیا ذره‌ای به نسل بعدی برود یا نه. هر ذره در جمعیت جدید با احتمال ۰/۰۲ همگرایی می‌باشد.

گام ۴ (توقف معیار)

اگر جمعیت جدید شرایط توقف را نداشته باشد، مراحل ۳-۴ به‌طور مستمر تا زمانی که مدل‌ها با حداقل خطای مدل برقرار شوند، انجام می‌شوند.

جمع‌آوری داده‌ها و پیش فرآیند رفتار سد

اکتساب داده‌ها

پارامترهای مورد استفاده شامل دمای آب، سطح آب، فشار دریچه، میزان رسوب‌گذاری، فشار منافذ، دمای هوا، حجم آب ورودی، مشخصات ویژه سد، شرایط بتن، سطح آب مخزن، تغییر مکان افقی و عمودی، اجزای اتصال انتقالی و شتاب زمین، قدرت، فشار، کشش و تنش بالا با استفاده از اطلاعات موجود در برگ مشخصات فنی و اطلاعات جمع‌آوری شده اعمال می‌شود (Bouaroudj *et al.*, 2019; Chen *et al.*, 2020). داده‌های استفاده شده در این تحقیق شامل ۳۶۵ روز، از ۱۳۹۷/۰۱/۳۱ تا ۱۳۹۸/۰۱/۳۱ به طول انجامیده است و همین‌طور فاصله نمونه‌گیری یک ماه می‌باشد. داده‌های رفتار سد برای پیش‌بینی رفتار آن به دو قسمت تقسیم شد: در وهله اول ۶۰۰ مجموعه داده تجهیزات سد برای آموزش مدل‌سازی الگوریتم بهینه‌سازی گروه ذرات مقدار حقیقی-رگرسیون لجستیک و آخرین ۱۲۰ مجموعه داده به‌عنوان داده‌های آزمایشی برای آنالیز عملکرد پیش‌بینی الگوریتم بهینه‌سازی گروه ذرات مقدار حقیقی-رگرسیون لجستیک استفاده شد.

پیش‌پردازش داده‌ها

در اصل، در حالت نرمال و با اشتراک‌گذاری داده‌ها، ورود داده‌های خام به الگوریتم باعث کاهش سرعت و دقت شبکه می‌شود. برای جلوگیری از چنین شرایطی و همچنین برای استانداردسازی مقادیر داده‌ها، داده‌های ورودی قبل از آموزش شبکه عصبی، باید استاندارد شوند، یعنی همه داده‌ها بین ۰ و ۱ برابر می‌شوند. پس از نرمال‌سازی، داده‌ها به دو دسته آموزشی و آزمایشی برای طراحی یک مدل شبکه تقسیم شدند. داده‌های آموزشی برای یافتن رابطه بین ورودی‌ها و خروجی‌های مشاهده‌ای توسط این مدل استفاده می‌شود.

ارزیابی و مقایسه عملکرد مدل پیشنهادی

محققان بسیاری آنالیزهای آماری مختلفی را برای ارزیابی روش‌های داده‌کاوی بکار گرفته‌اند. این‌گونه تحلیل‌ها معیار مناسبی را برای رتبه‌بندی روش‌های مختلف بر اساس دقت پیش‌بینی‌شان فراهم می‌آورند. برای ارزیابی و مقایسه مدل‌های رگرسیون لجستیک اسمی، شبکه عصبی و شبکه عصبی- فازی پیشنهاد شده، پارامترهای آماری مختلفی بکار گرفته شده‌اند. در این مطالعه نیز برای ارزیابی عملکرد روش‌های بکار گرفته شده از معیارهای ارزیابی ضریب تعیین (R^2)، ریشه میانگین مربعات خطا (RMSE)، ضریب پراکندگی (SI) و میانگین خطاهای انحرافی (MBE) استفاده گردید.

ضریب تعیین[1] (R^2): معیاری برای تعیین همبستگی نسبی بین دو مجموعه متغیر می‌باشد. این پارامتر در مرحله آموزش مدل، نقش مهم‌تری نسبت به مرحله آزمون ایفا می‌کند. ضریب تعیین، معیاری است از این که خط رگرسیون، چقدر خوب خوانده‌ها را معرفی می‌کند. اگر خط رگرسیون از تمام نقاط بگذرد توانائی معرفی همه متغیرها را دارد و هرچه از نقاط دورتر باشد نشان دهنده توانائی کمتر است. این معیار از (رابطه ۱) قابل محاسبه است.

$$R^2 = \left(1 - \frac{\sum_{i=1}^{n}(M_i - P_i)^2}{\sum_{i=1}^{n}(P_i - \bar{P}_i)^2}\right) \quad (۱)$$

ریشه میانگین مربعات خطا[2] (RMSE): برای اولین بار در سال ۱۹۸۰ توسط استیجر و لیند بوجود آمد که مقدار آن شدیداً به تعداد پارامترهای برآورد شده در مدل بستگی دارد. RMSE مشهورترین معیار خطا می‌باشد که خطاهای بزرگ را بیشتر از خطاهای کوچک جذب می‌کند (رابطه ۲).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_i - M_i)^2}{n}} \quad (۲)$$

ضریب پراکندگی[3] (SI): معرف پراکندگی متغیر به‌صورت مطلق می‌باشد، در واقع میزان پراکندگی داده‌ها از خط بهینه می‌باشد، اگر تمام داده‌ها درست پیش‌بینی شوند میزان پراکندگی صفر می‌باشد (رابطه ۳).

$$SI = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(Pi - Mi)^2}}{\bar{M}} \quad (۳)$$

میانگین خطاهای انحرافی[4] (MBE): از دیگر پارامترهای آماری مهم می‌باشد که هرچه به عدد صفر نزدیک باشد بهتر است، مقادیر مثبت برای MBE نشان از پیش‌بینی دست بالا و مقادیر منفی نشان از مقادیر دست پایین می‌باشد (رابطه ٤).

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(P_i - M_i) \quad (٤)$$

نتایج و بحث

در این مطالعه جهت آموزش مدل، داده‌ها به دو دسته آموزش و آزمون تقسیم شدند. همچنین جهت شبیه‌سازی مدل با داده‌های آزمون، داده‌ها بر اساس مشخصات داده‌های آموزش نرمال شدند. به همین منظور چندین ترکیب تصادفی برای تفکیک مجموعه‌های آموزش و آزمون انتخاب گردید تا نهایتاً دو مجموعه که از نظر

[1] Coefficient Of Determination

[2] Root Mean Square Error
[3] Scatter Index
[4] Mean Bias Error

آماری با هم سازگاری هستند، به دست آیند. با توجه به مطالعات انجام شده تقسیم داده‌ها به آموزش و آزمون به دو صورت ۸۰ به ۲۰ و ۷۰ به ۳۰ صورت می‌پذیرد و انتخاب هر یک از این روش بستگی به تعداد داده‌ها و ورودی‌ها دارد که در این مطالعه، ۸۰٪ از مجموعه داده برای آموزش (۱۳٤۷ داده) و ۲۰٪ از مجموعه داده (۳۳۷ داده) برای آزمون مدل پیشنهادی استفاده شده است. با توجه به فراوانی داده‌ها، جهت آموزش مدل‌ها سعی شد داده‌های آموزش به دسته‌های ۲۰۰، ٤۰۰، ٦۰۰، ۸۰۰، ۱۰۰۰ و ۱۳٤۷ تایی تقسیم و مرحله آموزش انجام شود تا بدین صورت حداقل مقدار داده که منجر به نتایج مناسب می‌شود، مشخص شود. جهت ارزیابی عملکرد مدل نیز از معیارهای آماری استفاده گردید که نتایج آن در (جدول ۲) آورده شده است.

جدول ۲. ارزیابی مدل‌های پیشنهاد شده

تعداد داده‌ها	R^2	RMSE(kPa)	Bias(kPa)	SI%
۲۰۰	۰/۷٥٥	۱۱/٤۷۳	۰/۰٤٥	۹/۹٦۷
٤۰۰	۰/۹۰٥	۱۲/۱۹۰	۰/۰۳۳	۹/۳٥۱
٦۰۰	۰/۹۲٦	۱۲/۷٦۳	۰/۰۲۲	۹/۳٦۹
۸۰۰	۰/۹۲٥	۱۳/٤۸۱	۰/۰۱٦	۸/۸۷۱
۱۰۰۰	۰/۹٦۲	۱٤/٦٥۷	۰/۰۱۱	۸/٤۷۲
۱۳٤۷	۰/۹٦٤	۱٤/۳٤۱	۰/۰۰٦	۷/۹۷٤

نتایج حاصل از مدل ورودی رگرسیون با توجه به معیار برازش برای ورود به انتخاب متغیرهای در (جدول ۳) قابل مشاهده است.

جدول ۳. برازش مدل‌های رگرسیونی ورودی الگوریتم ژنتیک

نسل	کروموزم	مدل رگرسیون	میزان برازش
۰	۰	لجستیک-اسمی	۰/۰۱٦
۰	Best	لجستیک-باینری	۰/۰۰٦
۱	Best	پروبیت-فضایی	۰/۰۰٦
۲	Best	ساده-خطی	۰/۰۰۸
۳	Best	چند متغیره	۰/۰۰۷
...	...	...	...
۹۷	Best	ساده-خطی	۰/۰۰٦
۹۸	Best	چند متغیره	۰/۰۰٦
۹۹	Best	لجستیک-اسمی	۰/۰۰٥

نتایج انتخاب پارامترهای مؤثر

به‌منظور تعیین مهم‌ترین عوامل مؤثر بر رفتار سد از آنالیز حساسیت در روش پیشنهادی استفاده گردید (جدول ٤). هر عامل و متغیری که بیشترین تأثیر در مدل پیشنهادی داشته باشد، حذف آن بر خطا و روند پیش‌بینی مدل تأثیر زیادی دارد. نتایج جدول ٤ بیانگر این است که متغیرهای دمای هوا، فشار منفذی و میزان رسوب بیشتر تأثیر در مدل و کمترین خطا را نشان داده‌اند و رگرسیون منتخب برای هر متغیر با توجه به شرایط حاکم و نوع داده متفاوت است و هیچ‌گونه برتری خاصی در میان رگرسیون‌های انتخاب شده در مدل وجود ندارد.

جدول ۴. آنالیز حساسیت مدل پیشنهادی در شناسایی متغیرهای مؤثر بر رفتار سد

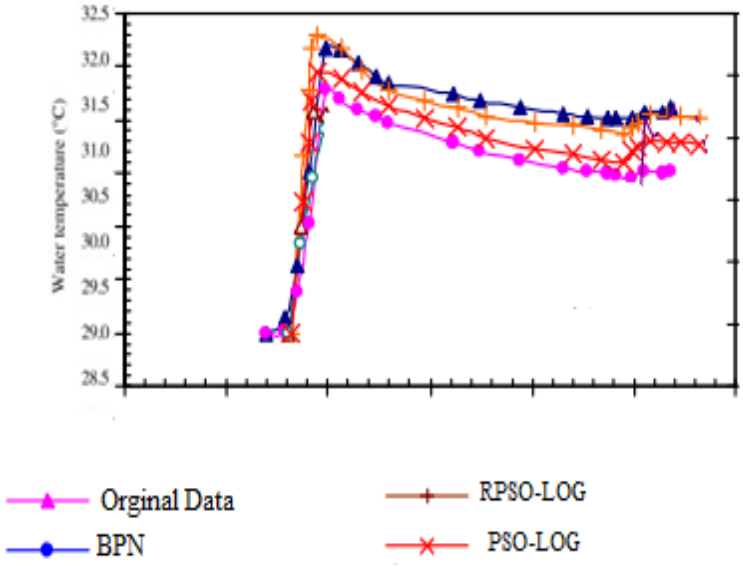
متغیر	رگرسیون منتخب	R^2	RMSE	MAE	درصد اهمیت
دمای آب	ساده	۰/۸۸	۰/۰۰۲	۰/۰۰۱	٥
سطح آب	ساده	۰/۸۹	۰/۰۰۱	۰/۰۰۱	٥
میزان رسوب	لوجیت	۰/۹۳	۰/۰۰۱	۰/۰۰۱	۱۲
فشار منفذی	ساده	۰/۹۷	۰/۰۰۱	۰/۰۰۱	۲۰
دمای هوا	لجستیک اسمی	۰/۹۸	۰/۰۰۱	۰/۰۰۱	۲۳
حجم آب ورودی	ساده	۰/۹۲	۰/۰۰۱	۰/۰۰۱	۱۰
مشخصات سد	لوجیت	۰/۸۸	۰/۰۰۲	۰/۰۰۱	۸
وضعیت بتن	پروبیت	۰/۹۰	۰/۰۰۲	۰/۰۰۱	۸
سطح آب مخزن	ساده	۰/۹۱	۰/۰۰۱	۰/۰۰۱	۹

طراحی مدل پیش‌بینی تجهیزاتی سد بر اساس الگوریتم بهینه‌سازی گروه ذرات مقدار حقیقی-رگرسیون لجستیک

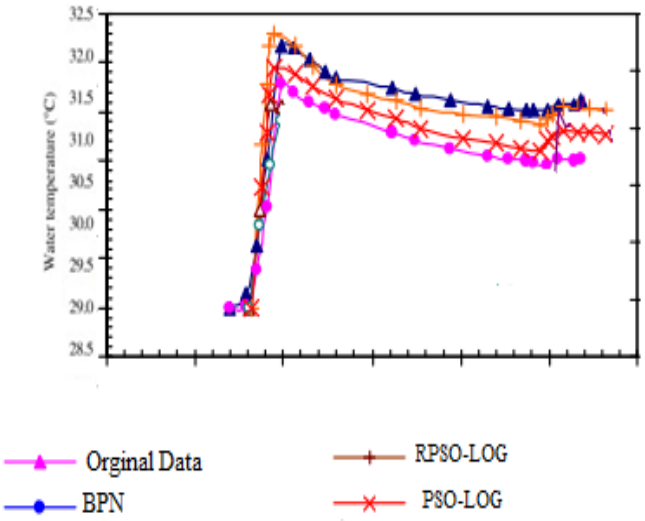
برای مدل رگرسیون لجستیک، هیچ فرآیند استانداردی برای تعیین پارامترهای آزاد وجود ندارد. در نتیجه، مدل پیشنهادی الگوریتم بهینه‌سازی گروه ذرات مقدار حقیقی-رگرسیون لجستیک به‌صورت دینامیکی مقادیر پارامترهای رگرسیون لجستیک را بهینه می‌کند. در وهله اول الگوریتم بهینه‌سازی گروه ذرات مقدار حقیقی و اعتبارسنجی متقابل پنج‌گانه برای جستجو استفاده شد و ترکیبات بهتری از پارامترهای رگرسیون لجستیک به دست آمد که ارزش اعتبارسنجی پنج‌گانه حداقل باشد. سپس، مدل پیش‌بینی تجهیزاتی سد با توجه به الگوریتم بهینه‌سازی گروه ذرات مقدار حقیقی - رگرسیون لجستیک ساخته می‌شود. سازه سیستم پیش‌بینی رفتار سد آب‌ورزی با توجه به الگوریتم بهینه‌سازی گروه ذرات مقدار حقیقی-رگرسیون لجستیک (در شکل ۴) نشان داده شده است. سیستم پیش‌بینی رفتار سد و تجهیزات آبی شامل اکتساب داده‌ها، پیش‌پردازش استانداردسازی داده‌ها، پیش‌بینی، نتایج آزمایش و کاربرد آن می‌باشد.

شکل ٤. طرح اولیه نقشه رابطه بین تحلیل فضایی و متغیرهای محلی

در طی آزمایش‌ها، از آنجایی که سطح محلول و دمای آب از مهم‌ترین عوامل تأثیرگذار بر مدل هستند، در این مطالعه دمای محلول آب و دما را تحت عنوان اهداف برای پیش‌بینی رفتار سد انتخاب شدند. برای پیش‌بینی رفتار سد، مقادیر تجهیزات امروزی نظارتی را به‌عنوان پارامترهای ورودی مدل الگوریتم بهینه‌سازی گروه ذرات مقدار حقیقی-رگرسیون لجستیک برای پیش‌بینی آنها به در نظر گرفتیم و سپس سطح کنترل و دمای آب مقایسه و آنالیز گردید. برای ارزیابی و مقایسه عملکرد روش ترکیبی



شکل ۵. پیش‌بینی مقدار سطح آب



شکل ٦. پیش‌بینی دمای آب

الگوریتم بهینه‌سازی گروه ذرات مقدار حقیقی-رگرسیون لجستیک، از روش عصبی مرسوم رگرسیون لجستیک و روش پس‌رو نیز برای مقایسه استفاده شد. برای شبکه عصبی پس‌رو، میزان یادگیری ۰/۰۸۶ و عملکرد فعال‌سازی سیگموئید می‌باشد. نتایج پیش‌بینی شده در زیر نشان داده شده است (شکل‌های ۵ و ۶).

شکل ۵ بیانگر نتایج پیش‌بینی شده از مقدار آب سطحی پیش‌بینی شده را به مدت نیم ساعت نشان می‌دهد. در این شکل محور قائم سطح آب بر حسب mg/L و محور افقی نمونه‌ای از آموزش زمان نمونه‌گیری می‌باشد. نتایج پیش‌بینی دمای آب به مدت نیم ساعت نیز در (شکل ۶) نشان داده شده است.

نتایج پارامترهای رفتار سد برای سه روش مختلف نشان داده شده است (جداول ۵ و ۶). بر اساس این جداول برای هر دو روش خطای ریشه میانگین مربع و خطای درصد میانگین مطلق، روش بهینه‌سازی گروه ذرات مقدار حقیقی- رگرسیون لجستیک توانایی تعمیم بهتر و پیش‌بینی فرآیند اعتبارسنجی را نسبت به روش‌های شبکه عصبی رگرسیون لجستیک و پس‌رو نشان می‌دهد. با استفاده از شاخص‌های رفتار تجهیزات سد، پیش‌بینی بهتر درجه حرارت آب و سپس سطح آب حاصل می‌شود. این امر بدان دلیل می‌باشد که دمای آب کمتر تحت تأثیر عوامل محیطی خارجی قرار می‌گیرد، درحالی‌که مقدار آب سطحی وابسته به مجموعه‌ای از فرآیندهای فعل و انفعالات فیزیکی-شیمیایی می‌باشد، در نتیجه تأثیرات سایر عوامل با گذشت زمان تغییر می‌کند. در کل، مدل پیش‌بینی بهینه‌سازی گروه ذرات مقدار حقیقی - رگرسیون لجستیک عملکرد بسیار خوبی دارد. این امر بدان دلیل

جدول ۵. مقایسه شاخص اندازه‌گیری عملکرد برای راهکارهای مختلف (مقدار سطح آب)

شبکه‌های عصبی پس‌رو مقدار آموزشی برابر با ۰/۰۸۸، سیگموئید			
خطای درصد میانگین مطلق ٪	۳/٦۸	۵/۹۱	۲۹/۹٦
خطای ریشه میانگین مربع	۰/۰۲۳٤	۰/۰۵۱۲	۰/۲۸۵

جدول ٦. مقایسه شاخص اندازه‌گیری عملکرد برای راهکارهای مختلف (دمای آب)

شبکه‌های عصبی پس‌رو مقدار آموزشی برابر با ۰/۰۸۸، سیگموئید			
خطای درصد میانگین مطلق ٪	۱/۸٤۵	٤/۸٤۷	۲۲/۲۷٦
خطای ریشه میانگین مربع	۰/۰۲۹۵	۰/۰۵۳٤	۰/۳۰۵

می‌باشد که بهینه‌سازی گروه ذرات مقدار حقیقی – رگرسیون لجستیک در عوض به حداقل رساندن ریسک تجربی که تعمیم عالی برای اندازه‌های نمونه کوچک را فراهم می‌سازد، اصل کاهش ریسک سازه‌ای را اجرا می‌کند. هیچ‌گونه توافقی بر سر توزیع مجموعه داده‌ها وجود ندارد. همچنین می‌تواند باعث تغییر در رویکرد داده‌ها شود؛ بنابراین، هنگامی که داده‌ها یک نوسان بزرگ از خود نشان می‌دهند، خطای پیش‌بینی بهینه‌سازی گروه ذرات مقدار حقیقی–رگرسیون لجستیک نیز کم می‌باشد.

نتیجه‌گیری

در این مطالعه از مدل‌های شبکه عصبی مبتنی بر رگرسیون با الگوریتم ژنتیک (GARANN) برای پیش‌بینی رفتار سد استفاده گردید. داده‌های رفتار سد برای پیش‌بینی رفتار آن به دو گروه طبقه‌بندی شدند که ابتدا ۶۰۰ مجموعه داده تجهیزات سد برای آموزش مدل‌سازی الگوریتم بهینه‌سازی گروه ذرات مقدار حقیقی– رگرسیون لجستیک و در ادامه ۱۲۰ مجموعه داده به‌عنوان داده‌های آزمایشی برای آنالیز عملکرد پیش‌بینی الگوریتم بهینه‌سازی گروه ذرات مقدار حقیقی – رگرسیون لجستیک بکار گرفته شد. به‌منظور ارزیابی عملکرد مدل استفاده شده از معیارهای آماری R2، RMSE، SI و MAB استفاده گردید. نتایج نشان داد که مدل استفاده در این تحقیق به‌خوبی آموزش دیده است و در پیش‌بینی فشار پیزومتریک در بدنه سد عملکرد قابل قبولی خواهد داشت. همچنین در معیار اول و دوم ارزیابی، نتایج حاصل از مدل شبکه عصبی ارائه شده با ضریب تعیین R^2= ۰/۹۳۰ و SI=۸/۵۸۷ همگرایی قابل قبولی نشان می‌دهد. نتایج مربوط به داده‌های آموزش مدل نیز بیانگر این است که میانگین (μ) و انحراف معیار (σ) مدل ارائه شده، برای داده‌های آموزش به ترتیب برابر با ۱/۳۴۱ و ۱/۵۲۶ می‌باشد و برای داده‌های صحت‌سنجی این مقادیر برابر با ۱/۵۷۶ و ۲/۲۴۷ می‌باشد که این نشان از عملکرد خوب مدل پیشنهادشده می‌باشد. در معیار احتمال تجمعی، مدل پیشنهاد شده با و P_{50}=۰/۹۴۰ و P_{90}=۱/۷۴۲ مبین این است که نتایج قابل قبول می‌باشد. نسبت مقادیر پیزومتریک پیش‌بینی‌شده به مقادیر قرائت‌شده برای حدود ۷۲ درصد داده‌ها در این مدل در حدود یک بوده، که بیانگر آموزش و قدرت پیش‌بینی مناسب این مدل می‌باشد. با توجه به جمع‌آوری مجموعه داده زیاد، مدل فوق به‌خوبی آموزش دید. بنابراین می‌توان این‌گونه نتیجه‌گیری کرد که الگوریتم بهینه‌سازی گروه ذرات مقدار حقیقی– رگرسیون لجستیک برای پیش‌بینی رفتار تجهیزاتی سد عملکرد بهتری از روش‌های شبکه عصبی پس‌رو و رگرسیون لجستیک دارد.

Reference:

Acerbi, L., & Ji, W. (2017). Practical Bayesian optimization for model fitting with Bayesian adaptive direct search. In: NIPS'17: Proceedings of the 31st International Conference on Neural Information. Processing Systems, 1834-1844.

Belmokre, A., Mihoubi, M. K., & Santillán, D. (2019). Analysis of dam behavior by statistical models: application of the random forest approach. KSCE Journal of Civil Engineering, 23(11), 4800- 4811.

Chambers, Lance D., ed. (2019). Practical handbook of particle swarm optimization algorithms: complex coding systems (Vol. 3). CRC press.

Chen, S., Gu, C., Lin, C., Zhang, K. & Zhu, Y. (2021). Multi-kernel optimized relevance vector machine for probabilistic prediction of concrete dam displacement. Engineering with Computers, 37(3), 1943-1959.

Colkesen, I., Sahin, E. K. & Kavzoglu, T. (2016). Susceptibility mapping of shallow landslides using kernel based Gaussian process, support vector machines and RPSO. Journal of African Earth Sciences, 118, 53-64.

Colkesen, I., Sahin, E. K., & Kavzoglu, T. (2016). Susceptibility mapping of shallow landslides using kernel-based Gaussian process, support vector machines, and RPSO. Journal of African Ea Sciences, 118, 53-64.

doi: 10.22065/jsce.2020.203517.1957. [in Persian]

Emami, S., Parsa, J., & Emami, H. (2021). Evaluating Cracks in Concrete Dams using Meta-heuristic Algorithms and Artificial Neural Networks. Journal of Structural and Construction Engineering, 8(Special Issue 2).

Feng, J., Liu, L., Wu, D., Li, G., Beer, M., & Gao, W. (2019). Dynamic reliability analysis using the extended support vector regression (X-SVM). Mechanical Systems and Signal Processing, 126, 368-391.

Gitoee, A., Faridi, A., & France, J. (2018). Mathematical models for response to amino acids: estimating the response of broiler chickens to branched-chain amino acids using support vector regression and neural network models. Neural Computing and Applications, 30(8), 2499-2508.

Gu, H., Yang, M., Gu, C., Cao, W., Huang, X. & Su, H. (2020). An analytical approach of behavior changes for concrete dam by panel data model. Steel and Composite Structures, 36(5), 521-531.

Han, H. G., Guo, Y. N., & Qiao, J. F. (2018). Nonlinear system modeling using a self-organizing recurrent radial basis function neural network. Applied Soft Computing, 71, 1105-1116.

Hussain, Z. F., Ibraheem, H. R., Alsajri, M., Ali, A. H., Ismail, M. A., Kasim, S., & Sutikno, T. (2020). A new model for iris data set classification based on linear support vector machine parameter optimization. International Journal of Electrical and Computer Engineering, 10(1), 1079-1084.

Kalita, D. J., Singh, V. P., & Kumar, V. (2020). A Survey on Logistic Regression Hyper-Parameters Optimization Techniques. In: Social Networking and Computational Intelligence, Springer, Singapore, 243-256.

Li, D., Wu, Y., Gao, E., Wang, G., Xu, Y., Zhong, H., & Wu, W. (2020). Simulation of Seawater Intrusion Area Using Feedforward Neural Network in Longkou, China. Water, 12(8), 2107. https://doi.org/10.3390/w12082107

Li, M., & Wang, J. (2019). An empirical comparison of multiple linear regression and artificial neural network for concrete dam deformation modeling. Mathematical Problems in Engineering. https://doi.org/10.1155/2019/7620948.

Ma, B., and Xia, Y. (2017). A tribe competition-based particle swarm optimization algorithm for feature selection in pattern classification. Applied Soft Computing, 58, 328-338.

Mata, J., Salazar, F., Barateiro, J., & Antunes, A. (2021). Validation of Machine Learning Models for Structural Dam Behaviour Interpretation and Prediction. Water, 13(19), 2717. https://doi.org/10.3390/w13192717

Nan, F., Ding, R., Nallapati, R., & Xiang, B. (2019). Topic Modeling with Wasserstein Autoencoders. arXiv preprint arXiv:1907.12374.

Neshat, M., Tabatabi, M., Zahmati, E., & Shirdel, M. (2016). A hybrid fuzzy knowledge-based system for forest fire risk forecasting. International Journal of Reasoning-based Intelligent Systems, 8(3-4), 132-154 [In Persian].

Oneto, L., Bisio, F., Cambria, E., & Anguita, D. (2016). Statistical learning theory and ELM for big social data analysis. IEEE Computational Intelligence Magazine, 11(3), 45-55. https://doi.org/10.1109/MCI.2016.2572540

Peng, K. L., Wu, C. H., & Goo, Y. J. (2004). The development of a new statistical technique for relating financial information to stock market returns. International Journal of Management, 21(4), 492-505.

Raj, J. S., & Ananthi, J. V. (2019). Recurrent neural networks and nonlinear prediction in support vector machines. Journal of Soft Computing Paradigm (JSCP), 1(01), 33-40. https://doi.org/10.36548/jscp.2019.1.004

Reis, L. P., de Souza, A. L., dos Reis, P. C. M., Mazzei, L., Soares, C. P. B., Torres, C. M. M. E., da Silva, L. F., Ruschel, A. R., Rêgo, L. J. S. & Leite, H. G. (2018). Estimation of mortality and survival of individual trees after harvesting wood using artificial neural networks in the amazon rain forest. Ecological Engineering, 112, 140-147.

Rong, M., Gong, D., & Gao, X. (2019). Feature selection and its use in big data: challenges, methods and trends. IEEE Access, 7, 19709-19725.

Samigulina, G., & Massimkanova, Z. (2019). Development of Smart-Technology for Forecasting Technical State of Equipment Based on Modified Particle Swarm Algorithms and Immune-Network Modeling. In: International Conference on Computational & Experimental Engineering and Sciences. Springer, Cham, 283-293.

Schratz, P., Muenchow, J., Iturritxa, E., Richter, J., & Brenning, A. (2018). Performance evaluation and hyperparameter tuning of statistical and machine-learning models using spatial data. arXiv preprint arXiv:1803.11266.

Shakarami, L., Javdanian, H., Sanayei, H. R. Z., & Shams, G. (2019). Numerical investigation of seismically induced crest settlement of earth dams. Modeling Earth Systems and Environment, 5(4), 1231-1238.

Siniscalchi, S. M., & Salerno, V. M. 2016. Adaptation to new microphones using artificial neural networks with trainable activation functions. IEEE transactions on neural networks and learn systems, 28(8), 1959-1965. https://doi.org/10.1109/TNNLS.2016.2550532.

Print ISSN: 2251-7480
Online ISSN: 2251-7400

Journal of
Water and Soil Resources Conservation (WSRCJ)

Web site:
https://wsrcj.srbiau.ac.ir

Email:
iauwsrcj@srbiau.ac.ir
iauwsrcj@gmail.com

Vol. 11
No. 4
Summer 2022

Received:
2021-09-29

Accepted:
2021-12-29

Pages: 115-127

Modeling The Behavior of Concrete Dams using Artificial Neural Network and Logistic Regression Methods

Fardin Saeid[1], Mohsen Irandoust[2*] and Navid Jalalkamali[3]

1) PhD Candidate, Department of Water Engineering, Kerman Branch, Islamic Azad University, Kerman, Iran.
2) Assistant Professor, Department of Water Engineering, Kerman Branch, Islamic Azad University, Kerman, Iran.
3) Assistant Professor, Department of Water Engineering, Kerman Branch, Islamic Azad University, Kerman, Iran.
*Corresponding Author email: mohsenirandost@yahoo.com

Abstract:

Background and Aim: Dam measurement and behavior assessment is a new issue that can be due to changes in available parameters to develop a model examining the behavior of individual parameters on the dam as well as on each other and analyze the changes and create the necessary policies. This study aims to propose a hybrid method involving logistic regression with particle swarm optimization algorithm with real value to predict the behavior of dam equipment.

Method: In this study, from 365 days data, from 04/20/2018 to 04/20/2019, of which 600 sets of dam equipment data including parameters of water temperature, water level, valve pressure, sedimentation rate, pore pressure, air temperature, inlet water volume, specific dam characteristics, concrete conditions, reservoir water level, horizontal and vertical displacement, transmission connection components and ground acceleration, strength, pressure, tensile and high stress were used for modeling. Real value-logistic regression and 120 datasets were used for modeling the should be added of particle group optimization algorithm. To evaluate the performance of the proposed method, four statistics including coefficient of determination (R^2), root mean square error (RMSE), scattering coefficient (SI), and means bias error (MBE) were used.

Findings: The results showed that the model has an acceptable performance in predicting piezometric pressure in the dam body. Also, the results of the artificial neural network model show acceptable convergence with R^2 = 0.930 and SSI = 8.587. The results related to the training data of the model also indicate that the mean (μ) and standard deviation (σ) of the proposed model are equal to 1.341 and 1.526 for the training data and these values for the validation data are equal to 1.576 and 2.247, respectively indicating the good performance of the proposed model. In the cumulative probability criterion, the proposed model with P_{50} = 0.940 and P_{90} = 1.742 indicates that the results are acceptable.

Results: The results indicate that the real value-logistic regression particle swarm optimization implements the principle of structural risk reduction instead of minimizing the experimental risk that provides excellent generalization for small sample sizes. The ratio of predicted piezometric values to read values for about 72% of the data in this model is about one, indicating the appropriate training and predictive power of this model. Finally, according to the evaluation criteria, the hybrid model performs better than the presented methods.

Keywords: Concrete dam, regression analysis, Particle swarm optimization algorithm, Sattarkhan Dam

10.30495/WSRCJ.2022.19230