

Assessment of Machine Learning Algorithms for Discharge Coefficient Prediction in Labyrinth-glory weirs: A Risk Analysis Approach

Hojatollah Safirzadeh¹ ,

Mohammad Heidarnejad² ,

Aslan Egdernezhad³ 

1. Former M.Sc. Student, Department of Civil Engineering, Ahvaz Branch, Islamic Azad University, Ahvaz, Iran.
2. *Corresponding author*, Associate Professor, Department of Water Science Engineering, Ahvaz Branch, Islamic Azad University, Ahvaz, Iran. mo_he3197@yahoo.com
3. Assistance Professor, Department of Water Science Engineering, Ahvaz Branch, Islamic Azad University, Ahvaz, Iran.

ABSTRACT

Morning glory spillways play a critical role in water flow management in dams and reservoirs, influenced significantly by the discharge coefficient. This coefficient determines the efficiency and risk of spillway performance under flood conditions. In this study, using 80 experimental datasets collected from two morning glory spillway inlet sections with square and circular zigzag shapes (featuring 4, 8, and 12 zigzags), two machine learning models—Support Vector Machine (SVM) and Gene Expression Programming (GEP)—were applied to simulate the discharge coefficient. Independent variables included the number of zigzags (n), Froude number (Fr), relative water head (H/P), and spillway shape index (R/D). Performance metrics (RMSE, MAE, R^2) were employed to evaluate the accuracy of the models. Among various SVM models, the RBF kernel with $\gamma = 0.1$ yielded the most optimal results. The training and testing phases for the circular spillway showed (RMSE, MAE, R^2) values of (0.9262, 0.0696, 0.0848) and (0.9820, 0.0346, 0.0398), respectively, while for the square spillway, these values were (0.9707, 0.073, 0.0904) and (0.9334, 0.0676, 0.0787). The GEP model demonstrated superior performance, particularly for the circular spillway with three genes, a head size of 9, and 45 chromosomes, yielding (RMSE, MAE, R^2) values of (0.9778, 0.0375, 0.0451) and (0.9811, 0.0315, 0.0396) in the training and testing phases, respectively. For the square section, the GEP model with 55 chromosomes achieved (RMSE, MAE, R^2) values of (0.9741, 0.0494, 0.0597) and (0.9591, 0.0503, 0.0594) for training and testing, respectively.

Keywords: Computational Fluid Dynamics, Diverted Flow, Data-Driven model, Performance Assessment.

ارزیابی عملکرد مدل‌های یادگیری ماشین در سرریزهای نیلوفری زیگزاگی بر مبنای تحلیل ریسک

حجت الله صفیرزاده^۱، محمد حیدرنژاد^۲، اصلا ن اگدرنژاد^۳

۱. دانش آموخته کارشناسی ارشد گروه عمران، واحد اهواز، دانشگاه آزاد اسلامی، اهواز، ایران.

۲. دانشیار گروه علوم و مهندسی آب، واحد اهواز، دانشگاه آزاد اسلامی، اهواز، ایران (mo_he3197@yahoo.com). نویسنده

مسئول

۳. استادیار گروه علوم و مهندسی آب، واحد اهواز، دانشگاه آزاد اسلامی، اهواز، ایران.

چکیده

سرریزهای نیلوفری، با تأثیرپذیری از ضریب دبی، در مدیریت جریان آب در سدها و مخازن نقش حیاتی دارند. ضریب دبی تعیین‌کننده کارایی و ریسک عملکرد آن‌ها در شرایط سیلابی است. در این راستا به کمک ۸۰ داده آزمایشگاهی گردآوری شده از دو مقطع ورودی سرریز نیلوفری با شکل‌های مربعی و دایروی زیگزاگی شده با تعداد چهار، هشت و دوازده عدد، از دو مدل ماشین بردار پشتیبان (SVM) و برنامه‌ریزی بیان ژن (GEP) برای شبیه‌سازی ضریب دبی استفاده شده است. تعداد زیگزاگ‌ها (n)، عدد فرود (Fr)، بار آبی نسبی (H/P) و شاخص شکل سرریز (R/D) به عنوان متغیرهای مستقل به کار گرفته شدند. شاخص‌های ارزیابی عملکرد (RMSE, MAE, R²) برای سنجش دقت خروجی مدل‌ها استفاده شدند. در بررسی مدل‌های مختلف SVM، تابع کرنل RBF با مقدار ۷ برابر ۱/۰ بهینه‌ترین نتایج را ارائه داد. مقادیر (RMSE, MAE, R²) در دوره‌های آموزش و آزمون برای این مدل به ترتیب (۰/۹۲۶۲، ۰/۰۶۹۶، ۰/۰۸۴۸) و (۰/۹۸۲۰، ۰/۰۳۴۶، ۰/۰۳۹۸) برای سرریز دایروی و (۰/۹۷۰۷، ۰/۰۷۳، ۰/۰۹۰۴) و (۰/۹۳۳۴، ۰/۰۶۷۶، ۰/۰۲۸۷) برای مقطع مربعی به دست آمد. در مدل GEP نتایج بهتری مشاهده شد، به گونه‌ای که مدل با سه ژن، اندازه هدا ۹ و ۴۵ کروموزوم، در سرریز دایروی با شاخص‌های (۰/۹۷۷۸، ۰/۰۳۷۵، ۰/۰۴۵۱) و (۰/۹۸۱۱، ۰/۰۳۱۵، ۰/۰۳۹۶) در مراحل آموزش و آزمون بهینه‌ترین عملکرد را داشت. برای مقطع مربعی، مدل با ۵۵ کروموزوم به ترتیب با مقادیر (۰/۹۷۴۱، ۰/۰۴۹۴، ۰/۰۵۹۷) و (۰/۹۵۹۱، ۰/۰۵۰۳، ۰/۰۵۹۴) در مراحل آموزش و آزمون ارزیابی شد.

کلمات کلیدی: ارزیابی عملکرد، تحلیل ریسک، سرریز نیلوفری، ضریب دبی، هوش مصنوعی.

مقدمه

ریسک در پروژه‌ها به عنوان رویدادهای نامعلوم و پیامدهای آنها بر اهداف پروژه توصیف می‌شود که می‌توانند بر زمان، هزینه و کیفیت پروژه تاثیرگذار باشند. در مواجهه با ریسکها، تجزیه و تحلیل، ارزیابی، کاهش و ارزیابی خسارات اقدامات مهم در مدیریت ریسک محسوب می‌شوند. این مفاهیم در زمینه‌های مهندسی، به‌خصوص در طراحی سدها و سازه‌های مرتبط، حائز اهمیت ویژه است. یکی از انواع سرریزها در سدها، سرریز نیلوفری است که در مواقعی که دیگر انواع سرریزها اقتصادی یا قابل اجرا نباشند، مورد استفاده قرار می‌گیرد. نکته مهم در این بین، اهمیت و لزوم توجه به تاج غیرخطی (طرح زیگزاگی تاج سرریز) در این نوع سرریزها می‌باشد. طرح زیگزاگی به دلیل افزایش طول تاج سرریز در مقایسه با طرح خطی، امکان عبور حجم بیشتری از جریان آب را در فضای محدود فراهم می‌کند. این نوع سرریزها به‌ویژه در شرایطی که محدودیت عرضی وجود دارد و نیاز به مدیریت سیلاب‌های با دبی بالا احساس می‌شود، گزینه‌ای مناسب و کارآمد هستند. (Faghih et al. (2009) در خصوص سد پیشین، در استان سیستان و بلوچستان، به ارزیابی و مقایسه تعدادی از روش‌های تحلیل کمی ریسک در برآورد سیلاب طراحی سرریز سدها پرداختند و سرریز شدن سیلاب از تاج سد را یکی از عوامل اصلی شکست انواع سدها بیان کردند. (Sharafati and Zahabiyou (2013) برآورد صحیح ریسک روگذری سد با در نظر گرفتن عدم قطعیت‌های متغیرهای تصادفی مؤثر بر آن را یکی از مباحث مهم در طراحی سدهای خاکی بر شمرند. (et Karamouz (2016) در پژوهش خود طراحی بهینه برای ابعاد سیستم انحراف آب در سدها با بررسی و آنالیز عدم قطعیت هیدرولیکی و ریسک هیدرولوژیکی را مورد بررسی قرار دادند و بیان داشتند احداث سیستم انحراف آب به دلیل قابل توجه بودن هزینه اجرای آن در سدها همواره طراحان را با مسائل و مشکلات فراوانی روبرو نموده است. (Maghrebi et al. (2018) به ارزیابی ریسک در اجزای سازه سرریز و اثرات پذیرش آن در عملکرد هیدرولیکی سازه در سد چندبر استان خراسان شمالی پرداختند. (Bahadori and Karimaei (2020) در پژوهش خود تعیین ارتفاع و ریسک روگذری سد های مخزنی بر مبنای تحلیل قابلیت اطمینان در سد نمرود را مورد بررسی قرار دادند. (Feyzi et al. (2020) با استفاده از روش تلفیقی FEMA و RAMCAP با رویکرد پدافند غیرعامل ارزیابی خطرپذیری سد بتنی برقابی لیرو را مورد بررسی قرار دادند و به این نتیجه رسیدند که امروزه با گسترش انواع تهدیدات انسان ساخت نظیر تهدیدات نظامی، تروریستی و خرابکارانه، توجه به کنترل و کاهش آسیب‌پذیری زیرساخت‌های حیاتی کشور به ویژه سامانه‌های تأمین آب مانند زیرساخت‌های سد و کسب آمادگی‌های لازم به منظور مقابله و بازیابی شرایط مطلوب در برابر انواع تهدیدات، امری ضروری است. (Eghbalizadeh et al. (2023) باز اندیشی در خصوص طراحی بخش‌های مختلف سامانه سرریز بر اساس تحلیل ریسک چند سطحی را در سد قزل‌داش در استان خراسان شمالی بررسی کردند. آنها با بررسی نظرات مدیران و دیدگاه‌های کارشناسان خبره در رابطه با موضوع روش ریسک چندسطحی، هر یک از فرایندها و بخش‌های اجرایی در حوزه پشتیبانی و رفع ریسک‌های ایمنی سد قزل‌داش را مورد تحلیل دوباره قرار دادند. (Rezapour Tabari and Hashempour (2018) توسعه مدل‌های هیبریدی-GWO-DSO و PSO-DSO را برای باز طراحی ابعاد بهینه سرریز کنگره‌ای مدنظر قرار دادند و بیان داشتند با توجه به اهمیت بازطراحی سرریز سدها برای ارائه ابعاد بهینه، رویکردی بر مبنای هیبرید الگوریتم‌های فراکاوشی ازدحام ذرات، گرگ‌های خاکستری و جستجوی مستقیم قابل پیشنهاد است. (Ebrahimzadeh et al. (2020) مدیریت ریسک روگذری سد حاجیلر چای استان آذربایجان شرقی را با شبیه سازی مونت کارلو و پویایی سیستم‌ها بررسی کردند. آنها بیان داشتند سدها در مهار سیلاب و کاهش خسارت آن نقش موثری دارند و از طرف دیگر در خلال سیلاب، بروز پدیده روگذری و عبور آب از روی سد همواره آن را تهدید می‌نماید. (Lucas et al. (2020) و (Frizell et al. (2020) بیان داشتند با توجه به رشد و گسترش احداث سدهای بزرگ و همچنین بالا رفتن استانداردهای ایمنی سدها، طراحی سرریزهای اقتصادی و مطمئن برای تخلیه سیلاب ورودی به مخزن سدها همواره یکی از مسائل مهم است. طبق نتایج به دست آمده توسط (Pfister et al. (2017) از چالش‌های مطرح سامانه سرریزها به خصوص سامانه سرریز آزاد وقوع ریسک های چند سطحی است که می‌تواند عملکرد و کارایی سرریزها و در نهایت کل سد را دچار مشکل کند. طبق نتایج به دست آمده

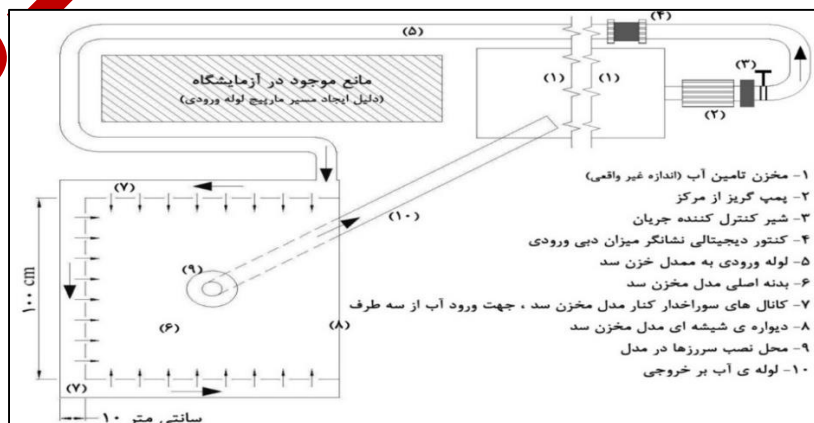
توسط (Borowski et al. 2020) از آنجایی که پدیده شکست سد ریسک و خطرپذیری را به همراه خواهد داشت بنابراین می‌توان از ارزیابی ریسک به عنوان ابزاری کارآمد و نو در عرصه مدیریت سدها به منظور کاهش خسارت‌ها استفاده نمود. (Gaagai et al. 2022) ریسک روگذری یک سد تاخیری را با در نظر گرفتن عدم قطعیت‌های ناشی از سیلاب ورودی به مخزن سد با در نظر گرفتن اثر روندیابی بررسی کردند. (Aouissi et al. 2021) با تحلیل اعتمادپذیری شرطی، ریسک روگذری فیتسوهی در آفریقای شمالی را بررسی کردند که شروط تحلیل اعتمادپذیری را مقدار دوره بازگشت سیل و تعداد دریاچه‌های غیرفعال برشمردند. (Lerma et al. 2014) با ارزیابی و مقایسه الگوریتم تکاملی اقدام به بهینه کردن طراحی قوانین عملیاتی در سیستم‌های منابع آب نمودند. نتایج پژوهش آنها نشان داد الگوریتم $SCE-UA$ برای رسیدن به راه حل بهینه احتیاج به تکرار کمتر و نتایج کارآمدتری نسبت به الگوریتم SSA دارد. (Sedighizadeh 2011) با استفاده روش بهینه‌سازی psa اقدام به بهینه‌سازی طول، قطر تونل و نوع نگهدارنده تونل و ارتفاع فرازبند نمودند. نتایج نشان داد که ابعاد بهینه و هزینه نهایی سیستم انحراف بستگی به دبی سیلاب دارند.

مرور پژوهش‌های انجام شده موید انجام تحقیقات مختلف در زمینه ارزیابی ریسک در قالب سناریوهای مختلف است که دارای مبانی هیدرولیکی و هیدرولوژیکی هستند. نکته حایز اهمیت پژوهش پیش رو استفاده از مدل‌های یادگیری ماشین شامل ماشین بردار پشتیبان و برنامه‌ریزی دین ژن براساس سناریوهای مختلف هیدرولیکی و فیزیکی برای ارزیابی مقدار ضریب دبی آبگذری سرریز نیلوفری با نصب زیگزاگ روی دهانه‌های ورودی است که در پژوهش‌های قبلی به آن پرداخته نشده است.

مواد و روش‌ها

داده‌های آزمایشگاهی

در این پژوهش از داده‌های گردآوری شده از مدل سرریز نیلوفری ساخته شده در آزمایشگاه هیدرولیک مجتمع عالی آموزشی و پژوهشی صنعت آب و برق خوزستان استفاده شده است. این مدل فیزیکی شامل سرریز نیلوفری با مقاطع دایره‌ای و مربعی شکل به ابعاد ۱۰۰ سانتی‌متر طول، ۱۰۰ سانتی‌متر عرض و ۵۰ سانتی‌متر ارتفاع است. جنس سه وجه جانبی و کف مدل از گالوانیزه با ضخامت ۲ میلی‌متر و یک وجه دیگر از شیشه با ضخامت ۱۰ میلی‌متر است. آبیگری از مخزن ذخیره آب توسط یک پمپ گریز از مرکز انجام می‌شود. در مسیر ورودی آب به فلوم، یک کنتور دیجیتال به منظور اندازه‌گیری دبی تعبیه شده است و ارتفاع آب روی سرریز به وسیله یک اشل قرائت می‌شود و آب خروجی از سرریز به صورت جریان آزاد توسط یک کانال به مخزن ذخیره آب باز گردانده می‌شود (شکل ۱). برای بررسی دقیق‌تر، مدل سرریز نیلوفری با توجه به پارامترهای مختلف هیدرولیکی مورد آزمایش قرار گرفت. سیستم آزمایشگاهی دارای امکاناتی برای تنظیم دقیق شرایط جریان و اندازه‌گیری پارامترهای مختلف بود. در این راستا، اندازه‌گیری ارتفاع آب روی سرریز و دبی خروجی به صورت دقیق انجام گرفت تا نتایج به دست آمده از آزمایش‌ها با مدل‌های تئوریک مقایسه شود.

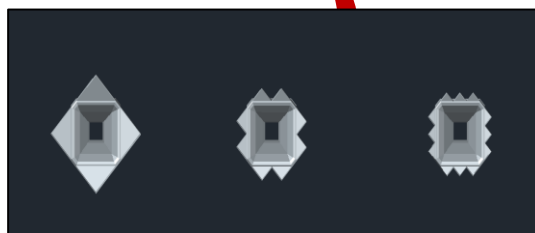


شکل ۱. نمایی کلی از پلان تجهیزات آزمایشگاهی

در پژوهش حاضر از سرریزهای دایره‌ای و مربعی زیگزاگی در مقام مقایسه با حالت بدون زیگزاگ (نمونه کنترلی) استفاده شده است (شکل ۲). سرریز دایره‌ای به ارتفاع ۱۸ سانتیمتر و قطر تاج ۲۰ سانتیمتر با مقیاس ۱:۱۴۰ از سرریز سد ویسکی تاون (WhiskeyTown) واقع در ایالت کالیفرنیا ایالات متحده آمریکا و از جنس پلی اتیلن طراحی و ساخته شد. ساخت مدل فیزیکی سرریز با مقطع مربع نیز بر پایه سرریز دایره‌ای انجام شد. در طراحی سرریز مربعی، از آنجا که هدف، رسیدن به مقطعی برای سرریز است که به لحاظ اجرا ساده‌تر، سریع‌تر و ارزان‌تر از مقطع معمول (دایره‌ای) باشد، نکات زیر مورد نظر قرار گرفتند: (۱) مقطع ورودی سرریز به جای دایره، مربع در نظر گرفته شد به طوری که طول ضلع مربع، معادل قطر تاج سرریز دایره‌ای (۲۰ سانتیمتر) لحاظ شد؛ (۲) در طراحی دیواره‌های سرریز مربعی هر قوس از دیواره سرریز دایره‌ای به چند شیب تقسیم شد به طوری که مجموعه شیبهای حاصل انطباق خوبی با قوس مرکب دیواره سرریز دایره‌ای داشته باشد. تعداد زیگزاگ‌های مورد استفاده برای هر مقطع چهار، هشت و دوازده عدد می‌باشد که نحوه اعمال آن روی سرریز مربعی در شکل ۳ نشان داده شده است. مقطعی از سرریز دایره‌ای با هشت زیگزاگ حین آزمایش در شکل ۴ نشان داده شده است.



شکل ۲. مقطع مربعی و دایره‌ای بدون استفاده از زیگزاگ



شکل ۳. مقطع مربعی با چهار، هشت و دوازده زیگزاگ



شکل ۴. سرریز نیلوفری با مقطع دایره با تعداد هشت زیگزاگ

در پژوهش حاضر از ۸۰ داده آزمایشگاهی مربوط به مدل آزمایشگاهی فوق‌الذکر به صورت ترکیبی مختلف از چهار پارامتر مستقل شامل عدد فرود (Fr)، تعداد زیگزاگ (n)، شعاع هیدرولیکی نسبی مجرای اصلی ($\frac{R}{D}$) و بار آبی نسبی ($\frac{H_0}{D}$) مورد بررسی قرار گرفتند.

برای انجام گام‌های آموزش و آزمون، ۷۰ درصد داده‌ها برای گام آموزش و ۳۰ درصد داده‌ها برای گام آزمون مورد استفاده قرار گرفتند. در خصوص مدل *SVM*، توابع کرنل مختلف برای این ترکیب داده‌ها با انتخاب مقادیر مناسب برای پارامتر تنظیمی γ به کمک آزمون و خطا بهینه‌سازی شدند. برای مدل *GEP* نیز به ازای مقادیر مختلف از پارامترهای تنظیمی و عملگرها ریاضی، بهینه‌ترین مدل انتخاب شدند. از نرم افزار متلب *MATLAB* برای اجرای این دو مدل استفاده شد.

تحلیل حساسیت

برای ارزیابی مقدار تاثیرپذیری ضریب دبی از هر یک از متغیرهای مستقل، آنالیز حساسیت با استفاده از آزمون گاما تست انجام شد. در این آزمون، که اهمیت هر یک از متغیرهای مستقل در تغییرات متغیر وابسته از طریق اعمال یا عدم اعمال هر یک از متغیرهای مستقل انجام می‌شود چهار پارامتر گاما، گرادیان، خطای استاندارد و پارامتر *V-Ratio* محاسبه می‌شود. کمترین مقدار به دست آمده برای این چهار پارامتر در هر ترکیب نشان از بهینه‌ترین ترکیب برای بررسی پارامترهای موثر در پارامتر وابسته است. نتیجه تحلیل حساسیت در جدول ۱ ارائه شده است. همانطور که مشخص است هر چهار پارامتر باید در بررسی تغییرات ضریب دبی جریان دخیل باشند زیرا ردیف اول جدول کمترین مقدار هر چهار پارامتر ورودی را دارد. با بررسی نقش هر یک از متغیرهای مستقل در میزان تغییرات پارامتر وابسته مشخص شد متغیرهای *Fr*، *H/P*، *R/D* و n به ترتیب بیشترین تا کمترین تاثیر را روی ضریب دبی جریان دارند.

جدول ۱. خلاصه نتایج تحلیل حساسیت

Test No.	Gamma	Gradient	Standard error	V-ratio	Mask
۱	۰/۰۰۸۷۱۱	۰/۹۱۹۰۲۷	۰/۰۰۵۳۹۸	۰/۲۰۷۲۸۸	۱۱۱۱
۲	۰/۰۱۳۳۴۴	۰/۱۸۳۰۴۱	۰/۰۰۵۵۹۳	۰/۳۱۷۵۳۵	۱۰۱۱
۳	۰/۰۱۳۳۶۸	۰/۲۶۴۳۳۲	۰/۰۰۶۰۴۱	۰/۳۱۸۰۹	۱۱۱۰
۴	۰/۰۲۴۲۱۱	۰/۷۳۲۹۴۶	۰/۰۰۴۶۵۵	۰/۵۷۶۱۰۲	۱۰۱۰
۵	۰/۰۲۷۸۰۸	۰/۱۶۱۲۴۸	۰/۰۰۵۳۴۹	۰/۶۶۱۶۹۱	۱۱۱۰
۶	۰/۰۲۹۷۸۳	۰/۰۰۰۷۸۶	۰/۰۱۴۱۴۵	۰/۷۰۸۶۸۵	۱۰۱۱
۷	۰/۰۲۷۵۵۵	۰/۲۲۰۰۲۶	۰/۰۰۵۴۱۱	۰/۶۵۵۶۷۳	۱۱۰۰
۸	۰/۰۲۷۰۰۹	۰/۳۰۱۱۰۶	۰/۰۰۲۳۵۵	۰/۶۴۴۶۱۲	۱۱۰۰
۹	۰/۰۱۶۹۸۳	۰/۵۸۹۱۴۲	۰/۰۰۲۱۰۳	۰/۶۴۲۰۵۸	۱۰۰۰
۱۰	۰/۰۲۹۴۶۶	۰/۰۰۱۱۷۶	۰/۰۱۴۰۵۷	۰/۷۰۱۱۵۷	۱۰۱۰
۱۱	۰/۰۲۶۹۴۴	۰/۲۷۸۱۹	۰/۰۰۳۷۰۹	۰/۶۴۱۰۴۴	۱۰۰۰
۱۲	۰/۰۳۰۰۲۵	۰/۰۰۰۷۲۵	۰/۰۱۵۱۶۸	۰/۷۱۴۴۴۲	۱۰۰۱
۱۳	۰/۰۴۹۱	-۰/۰۰۰۰۲۶	۰/۰۱۵۶۴	۱/۱۶۸۳۵۶	۱۱۰۰
۱۴	۰/۰۴۲۰۲۵	.	.	۱	۱۰۰۰
۱۵	۰/۰۵۷۳۴۹	-۰/۰۰۰۰۳۹	۰/۰۱۴۱۲۵	۱/۳۶۴۶۳۷	۱۰۰۰

مروری بر مدل‌های یادگیری ماشین

مدل SVM

ماشین بردار پشتیبان یک الگوریتم یادگیری ماشین نظارت شده قدرتمند است که برای طبقه‌بندی و رگرسیون استفاده می‌شود. مدل SVM با یافتن بهترین مرز تصمیم (*hyperplane*) بین داده‌های دو دسته مختلف، آن‌ها را از هم جدا می‌کند. این مدل با یافتن حاشیه (*margin*) بیشینه بین دو دسته، بهترین مرز تصمیم را پیدا می‌کند. این حاشیه توسط بردارهای پشتیبان (*support vectors*) تعریف می‌شود که نزدیک‌ترین نقاط به مرز تصمیم هستند. مرز تصمیم SVM توسط یک تابع خطی به شکل زیر تعریف می‌شود:

$$W^T.X+b=0 \quad (۱)$$

که در آن W بردار وزن، X بردار ویژگی و b بایاس می‌باشد. هدف اصلی مدل SVM یافتن W و b به گونه‌ای است که حاشیه بیشینه شود. به عبارتی هدف بیشینه کردن مقدار $\|W\|$ با قید زیر است:

$$y_i(W^T.X_i+b) \geq 1 \quad i=1, 2, \dots, N \quad (۲)$$

که در آن y_i برچسب کلاس داده i ام و X_i بردار ویژگی داده i ام است. برای حل این مسئله بهینه‌سازی، از روش لاگرانژ استفاده می‌شود. تابع لاگرانژ به صورت زیر تعریف می‌شود:

$$L(W,b,\alpha) = \frac{1}{2} \|W\|^2 - \alpha \sum_{i=1}^N \alpha_i (y_i(W^T.X_i+b)-1) \quad (۳)$$

با حل این مسئله بهینه‌سازی، می‌توان W و b را به دست آورد. در بسیاری از موارد، داده‌ها به صورت خطی قابل جداسازی نیستند. در این موارد، از تابع هسته (ترفند کرنل) استفاده می‌شود. تابع هسته یک تابع غیرخطی است که داده‌ها را به یک فضای با ابعاد بالاتر نگاشت می‌کند و در این فضا، داده‌ها به صورت خطی قابل جداسازی می‌شوند (Vapnik, 1995). انواع تابع کرنل در مدل SVM در جدول ارائه شده‌اند (Fuladipana and Majedi-Asl, 2022).

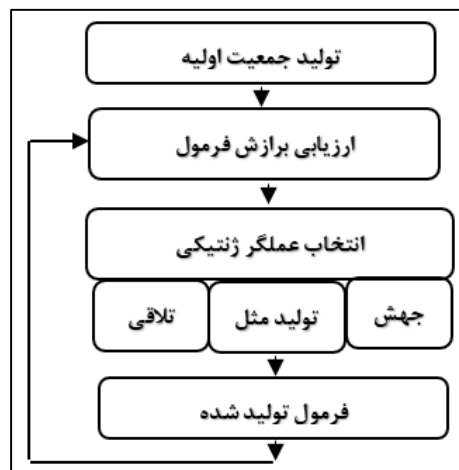
جدول ۲. انواع توابع کرنل

نام کرنل	رابطه‌ی کرنل
خطی	$K(x_i, x_j) = (x_i, x_j)$
چند جمله‌ای	$K(x_i, x_j) = [(x_i, x_j) + 1]^d$
تابع پایه شعاعی	$K(x_i, x_j) = \exp[-\frac{\ x_i - x_j\ ^2}{2\sigma^2}]$
سیگموئید	$K(x_i, x_j) = \tanh[-\alpha(x_i, x_j) + c]$

مدل GEP

الگوریتم ژنتیک، که توسط Ferreira (2001) بر اساس نظریه داروین ارائه شد، پایه و اساس مدل برنامه‌ریزی بیان ژن است. این الگوریتم، جمعیتی از افراد را بر اساس شاخص برازش انتخاب می‌کند و با استفاده از عملگرهای ژنتیکی، تغییرات ژنتیکی را اعمال می‌نماید. مدل GEP از این الگوریتم ژنتیکی برای ایجاد جمعیت اولیه‌ای از فرمول‌ها استفاده می‌کند که از ترکیب تصادفی توابع ریاضی، متغیرهای مسئله و اعداد ثابت به دست می‌آیند. سپس، افراد این جمعیت بر اساس شاخص‌های برازش موجود در مدل، ارزیابی می‌شوند. در مرحله بعد، جمعیت جدیدی از فرمول‌ها تولید می‌شود و این فرایند تا رسیدن به بیشینه مقدار تولید، تکرار می‌گردد. شکل

۲، الگوریتم استفاده شده در مدل *GEP* را نشان می‌دهد که در آن، چهار مرحله اصلی شامل ایجاد جمعیت اولیه، ارزیابی افراد، ایجاد جمعیت جدید و تکرار فرایند تا رسیدن به نتیجه بهینه، به تصویر کشیده شده است.



شکل ۵. فلوچارت الگوریتم برنامه‌ریزی بیان ژن

مقایسه کارکرد مدل‌ها

در این پژوهش، مدل‌ها در طی فرآیندهای آموزش و آزمون ارزیابی شدند تا بهترین تنظیمات پارامترها به دست آید. این مرحله مستلزم استفاده از معیارهای ارزیابی عملکرد است. برای این منظور، سه معیار شامل ریشه میانگین مربع خطاها (*RMSE*)، میانگین قدر مطلق خطاها (*MAE*) و ضریب تعیین، با استفاده از فرمول‌های زیر به کار گرفته شدند:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (۴)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (۵)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (۶)$$

که در آنها y_i مقادیر مشاهداتی و $\hat{y}_i$ مقادیر محاسباتی از داده i ام، n نشان دهنده تعداد کل داده‌ها و $\bar{y}$ بیانگر مقدار میانگین داده‌های مشاهداتی هستند. خطای جذر میانگین مربعات (*RMSE*)، ابزاری مناسب برای مقایسه میان مقدار پیش‌بینی شده توسط مدل و مقدار واقعی با بعد موجود است. همچنین هر چقدر شاخص به صفر نزدیک باشد، مدل از دقت بالایی برخوردار است. ضریب تبیین یکی از مهم‌ترین معیارها در تحلیل رگرسیون است و بیانگر مقدار تغییرات متغیر وابسته که توسط متغیرهای مستقل توضیح داده می‌شود. این ضریب بین ۰ و ۱ قرار دارد و به صورت درصدی از توضیح واریانس متغیر وابسته توسط مدل رگرسیون تعبیر می‌شود. مقدار صفر بیانگر این است که مدل هیچ توضیحی برای تغییرات متغیر وابسته ندارد. مقدار یک نشان می‌دهد مدل تمام تغییرات متغیر وابسته را به طور کامل توضیح می‌دهد. مقدار بین صفر و یک بیانگر آن است که بخشی از تغییرات متغیر وابسته توسط مدل توضیح داده می‌شود. ضریب تبیین بالا نشان‌دهنده مدل خوبی است که توانسته است تغییرات متغیر وابسته را به خوبی توضیح دهد. با این حال، باید توجه داشت که به تنهایی نمی‌تواند تضمین‌کننده کیفیت مدل باشد و ممکن است مدل با بالا همچنان دچار مشکلاتی مثل همبستگی کاذب یا بیش‌برازش (*overfitting*) باشد. ضریب میانگین خطای مطلق (*MAE*) یکی از معیارهای ارزیابی عملکرد مدل‌های پیش‌بینی و رگرسیون است. این ضریب بیانگر میانگین قدر مطلق اختلافات بین مقادیر واقعی و مقادیر پیش‌بینی شده توسط مدل است. به عبارت

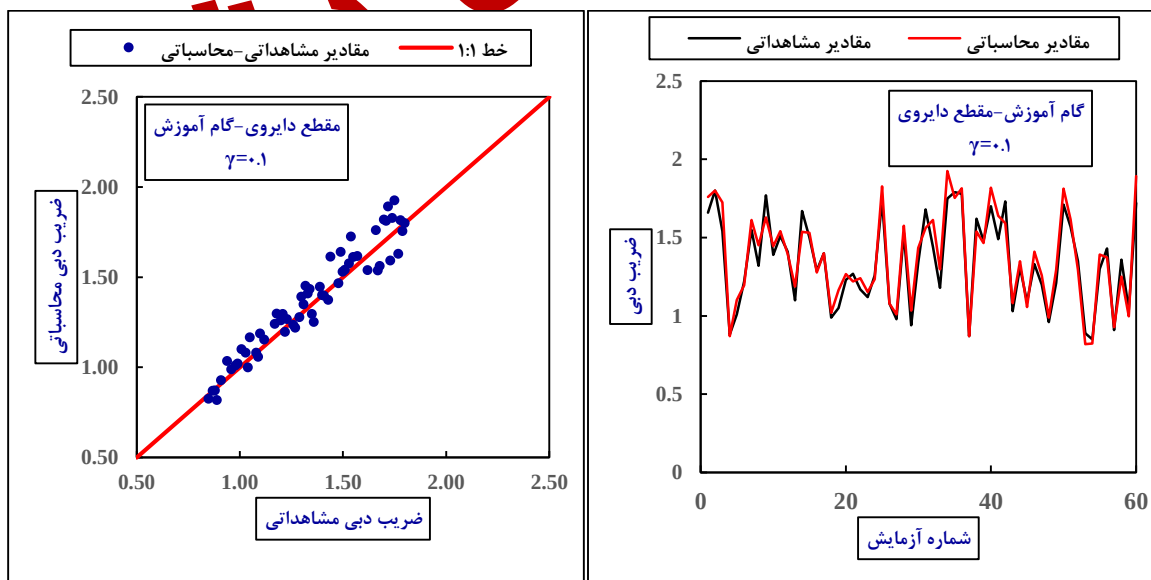
دیگر، MAE نشان می‌دهد که مدل به طور متوسط چه مقدار خطا دارد. مقدار صفر نشان می‌دهد مدل به طور کامل و بدون خطا تمامی مقادیر واقعی را پیش‌بینی کرده است. هر چه MAE بزرگ‌تر باشد، مدل خطای بیشتری در پیش‌بینی دارد. سادگی در محاسبه و تفسیر و عدم تأثیرپذیری از مجذور خطاها از مزایای این شاخص آماری است. بنابراین می‌توان از MAE برای مقایسه عملکرد مدل‌های مختلف استفاده کرد و مدلی که MAE کمتری دارد، عملکرد بهتری دارد (Fuladipannah et al., 2020).

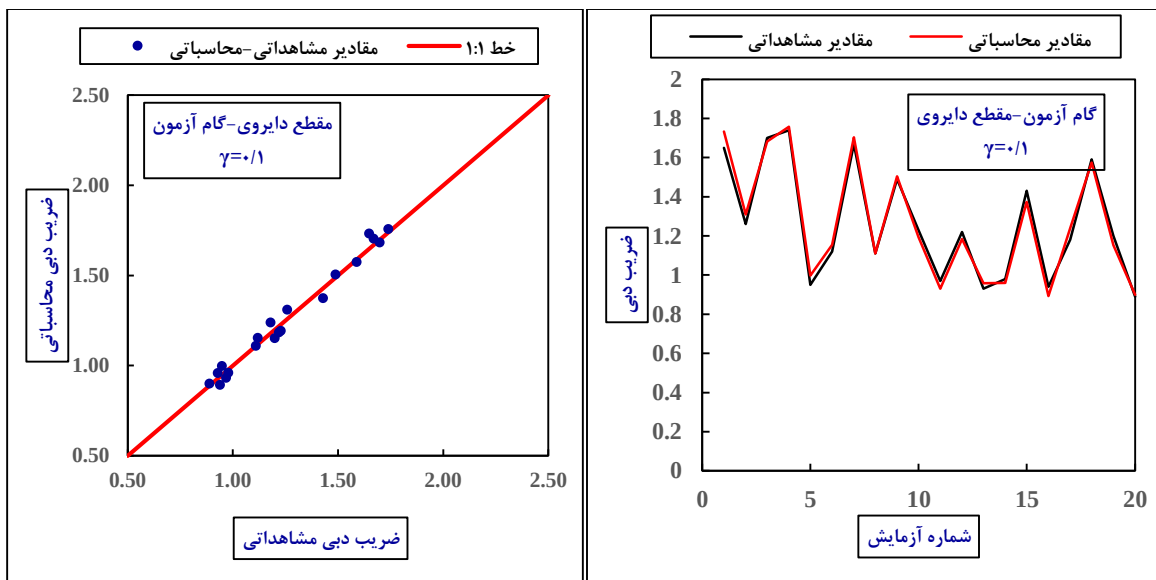
بحث و نتایج

با استفاده از داده‌های آزمایشگاهی، دو مدل یادگیری ماشین SVM و GEP روی ۸۰ داده مربوط به ضریب دبی سرریزهای نیلوفری زیگزاکی اجرا شد. مهم گام‌های آموزش و آزمون از این تعداد داده به ترتیب ۷۰ و ۳۰ درصد لحاظ شد. با استفاده از فرآیند آزمون خطا، مقادیر مختلفی برای پارامترهای تنظیمی مدل SVM شامل C ، ϵ و γ برای هر یک از توابع کرنل در نظر گرفته شدند و در نهایت مقادیر شاخص‌های ارزیابی عملکرد ($RMSE$ ، MAE ، R^2) نشان دادند تابع RBF بهینه‌ترین خروجی را به ازای مقادیر $C=65$ ، $\epsilon=0.18$ و $\gamma=0.1$ برای هر دو مقطع دایروی و مربعی دارد. مقدار ضرایب ارزیابی عملکرد مدل SVM برای دو مقطع دایروی و مربعی در جدول ۳ طی گام‌های آموزش و آزمون نشان داده شده است. بیان گرافیکی از میزان انطباق‌پذیری ضرایب دبی مشاهداتی و محاسباتی مقاطع دایروی و مربعی به ترتیب در شکل ۶ و شکل ۷ نشان داده‌اند.

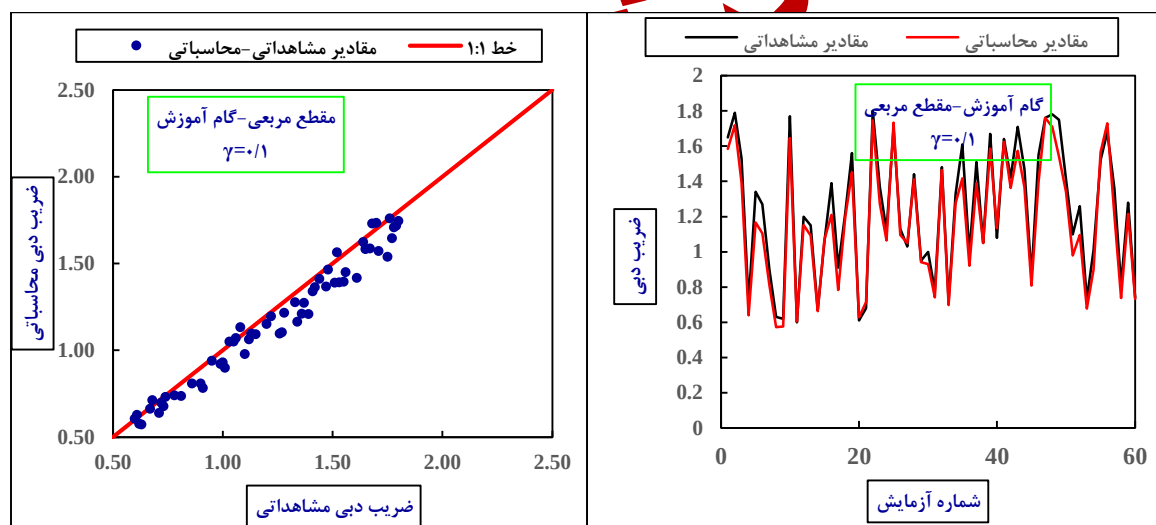
جدول ۳. خلاصه نتایج ارزیابی عملکرد مدل SVM

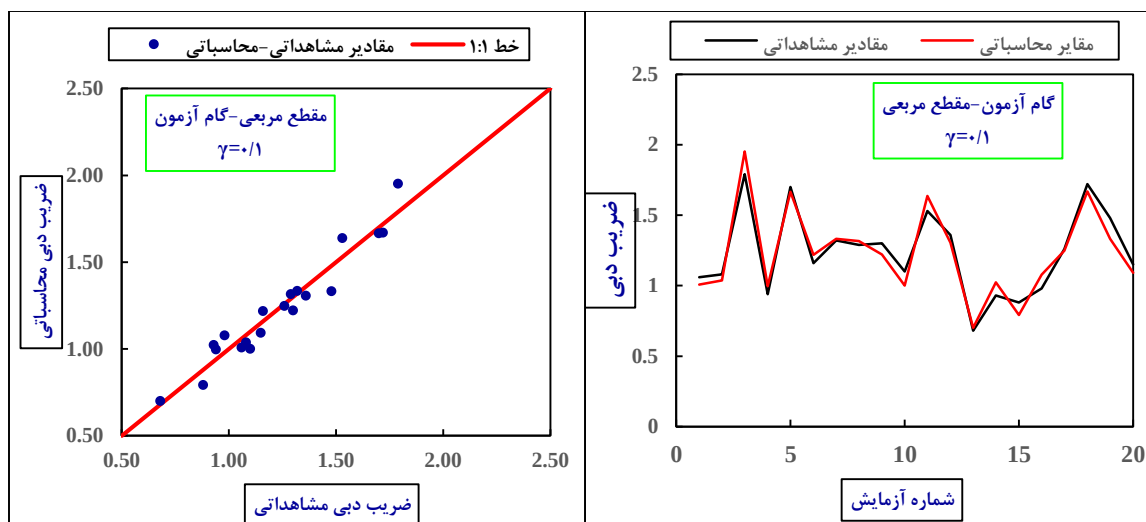
مجرای دایروی						
ردیف	مقدار γ	گام آموزش			گام آزمون	
		R^2	MAE	$RMSE$	R^2	MAE
۱	۰/۱	۰/۰۸۴۸	۰/۰۶۹۱	۰/۰۹۲۶۲	۰/۰۳۹۸	۰/۰۳۴۶
مجرای مربعی						
ردیف	مقدار γ	گام آموزش			گام آزمون	
		R^2	MAE	$RMSE$	R^2	MAE
۱	۰/۱	۰/۰۹۰۴	۰/۰۷۳۰	۰/۰۹۷۰۷	۰/۰۷۸۷	۰/۰۶۷۶





شکل ۶ عملکرد مقدار بهینه ضریب γ برای تابع کرنل RBF در سرریز دایروی



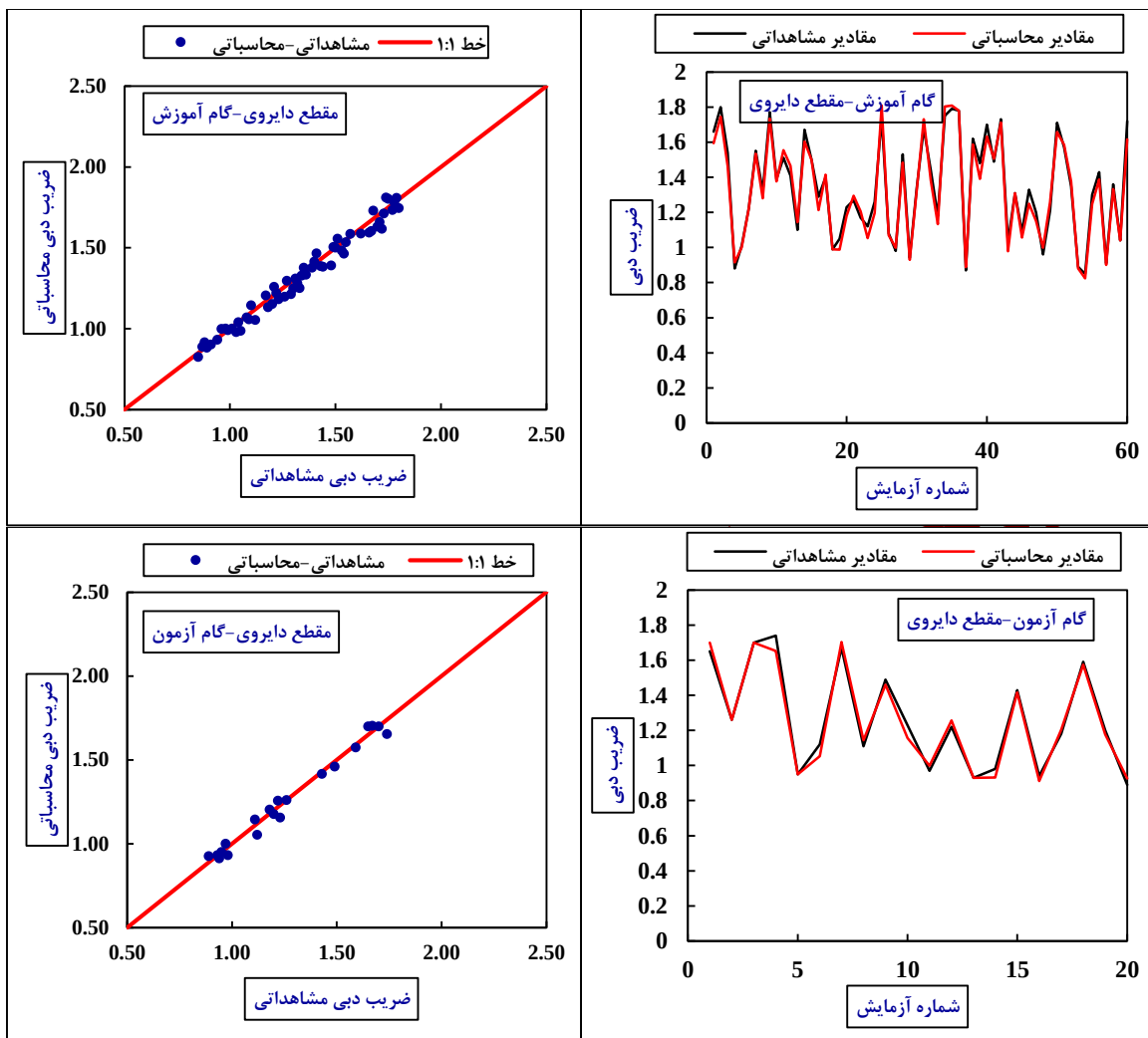


شکل ۷. عملکرد مقدار بهینه ضریب γ برای تابع کرنل RBF در سرریز مربعی

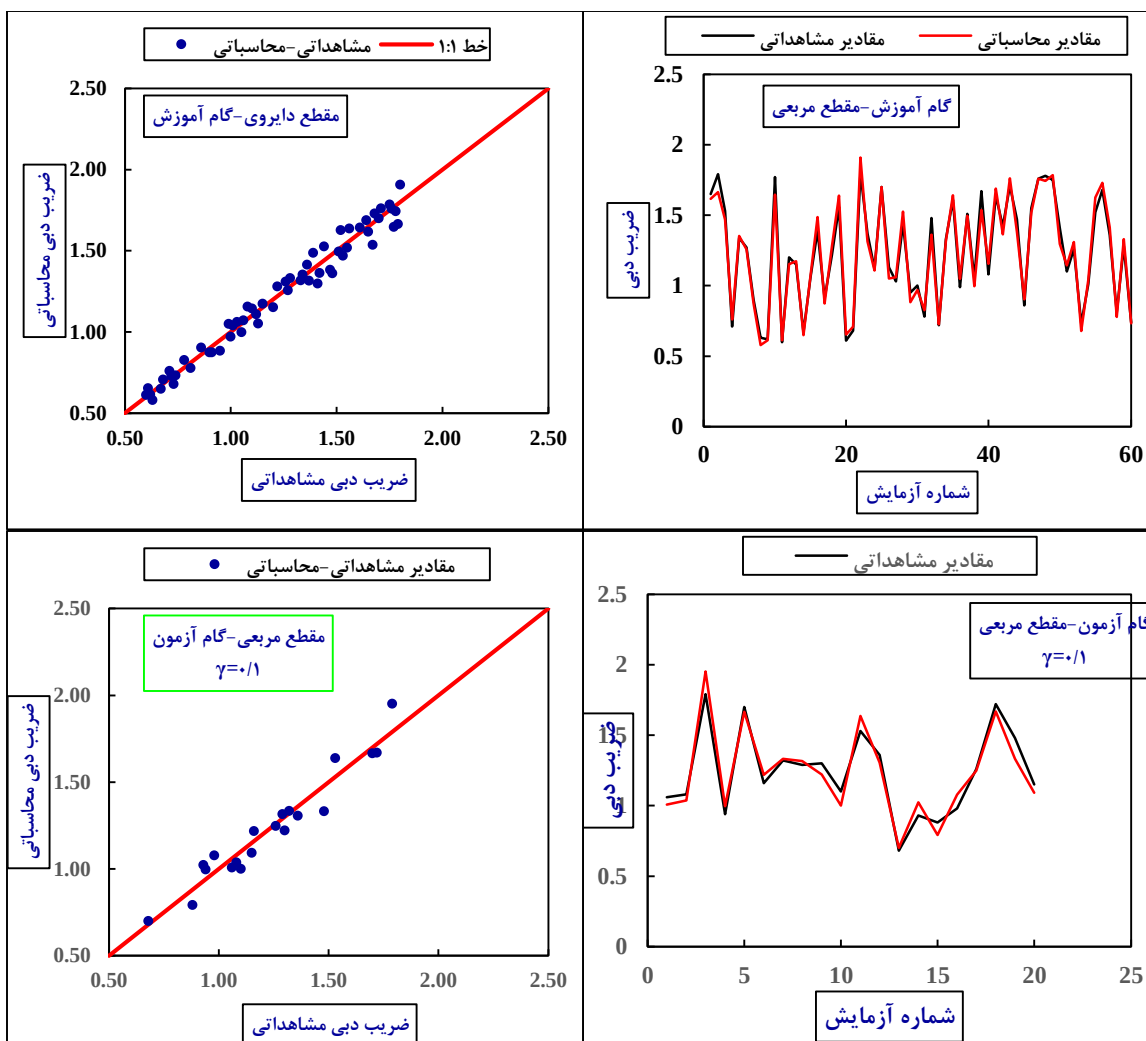
در اجرای مدل GEP نیز بررسی گزینه‌های مختلف برای رسیدن به بهینه‌ترین خروجی با استفاده از روش آزمون و خطا بر روی سه پارامتر تنظیمی اندازه سر، تعداد ژن و تعداد کروموزم انجام گرفت. خلاصه محاسبات مربوط به بهینه‌ترین ترکیب برای مدل GEP در جدول ۲ ارائه شده است. میزان انطباق‌پذیری داده‌های مشاهداتی و محاسباتی ضریب دبی در قالب بیان گرافیکی در دو مقطع دایروی و مربعی به ترتیب در شکل ۸ و شکل ۹ ارائه شده‌اند.

جدول ۲. ارزیابی عملکرد بهینه‌ترین ترکیب مدل GEP

پارامتر	مقدار	
	سرریز دایره‌ای	سرریز مربعی
تابع برازندگی	RMSE	RMSE
تابع اتصال ژن‌ها	+	
تعداد کروموزم	۴۵	۵۵
تعداد ژن	۳	۳
اندازه رأس	۹	۹
نرخ ترکیب	۰/۱	۰/۱
نرخ جهش	۰/۱	۰/۱



شکل ۸. بیان گرافیکی از مدل بهینه GEP در سرریز دایروی



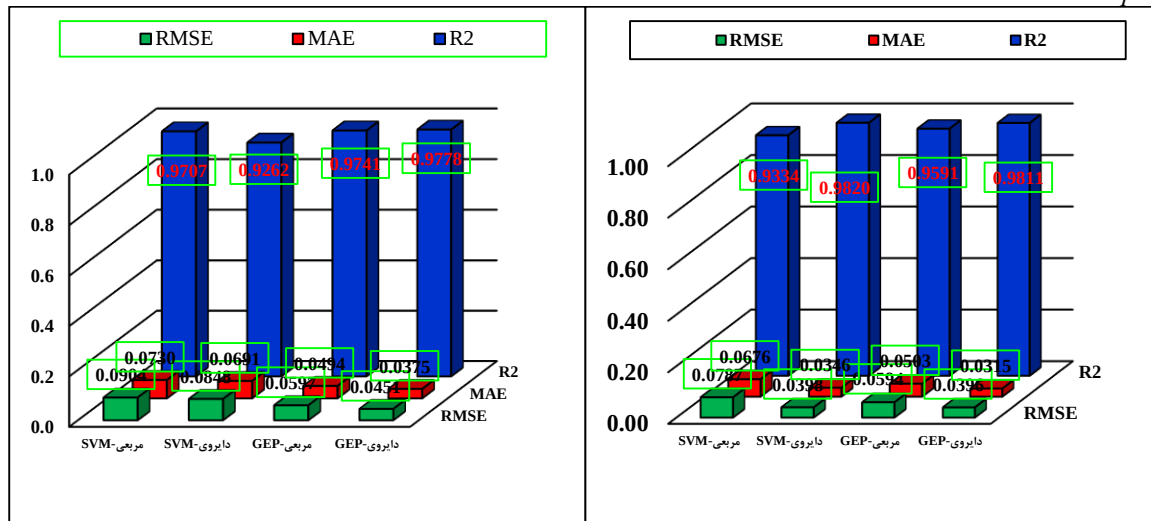
شکل ۹. بیان گرافیکی از مدل بهینه GEP در سرریز مربعی

حال با توجه به مقادیر به دست آمده برای شاخص‌های ارزیابی عملکرد، لازم است بهترین مدل از بین دو مدل GEP و SVM برای شبیه سازی ضریب دبی سرریز نیلوفری-زیگزاکی انجام شود. نمودارهای ارائه شده در شکل ۱۱ مقایسه بین بهترین خروجی مدل‌های GEP و SVM را در دو گام آموزش و آزمون نشان می‌دهند. با توجه به نتایج به دست آمده می‌توان نتیجه گرفت مدل GEP نسبت به مدل SVM از دقت بیشتری در پیش‌بینی ضریب دبی برخوردار است. به همین دلیل، در ادامه برای تحلیل شرایط مختلف مقادیر پارامترهای هیدرولیکی و هندسی روی ضریب دبی جریان از مدل GEP استفاده شده است. خروجی مدل GEP در قالب بیان درختی در شکل ۱۱ ارائه شده است. همانطور که مشخص است توابع مورد استفاده در این مدل شامل $Sqrt$ ، $-$ ، $\times$ ، $+$ ، $/$ ، $\sin$ ، $\cos$ ، X^2 و X^3 هستند. معادله ریاضی متناظر با بیان درختی ارائه شده در شکل ۱۱ به قرار زیر است:

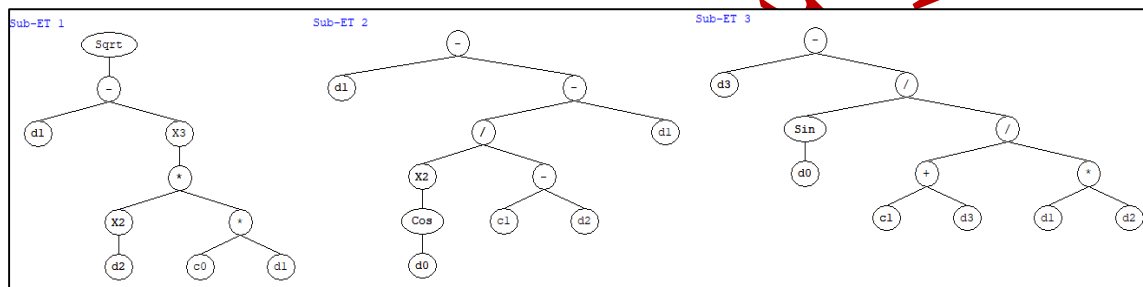
$$C_d = \sqrt{\frac{R}{D} - (Fr^2 \times (-6.151184 \frac{R}{D})^3 + \left(\frac{R}{D} - \left(\frac{\cos^2 n}{4.580292 - \frac{R}{D}} - \frac{R}{D} \right) \right) + \left(\frac{H}{P} - \frac{\sin n}{-1.315918 + \frac{H}{P}} \right) \left(\frac{R}{D} \times Fr \right)} \quad (7)$$

تغییرات ضریب دبی در برابر متغیرهای مستقل در **Error! Reference source not found.** ارائه شده است. هر محدوده رنگی، بیانگر دامنه تغییرات ضریب دبی سرریز است. این شکل نحوه تغییرات ضریب دبی را به صورت تابعی از $(\frac{R}{D}, \frac{H}{P})$ ، (Fr, n) و

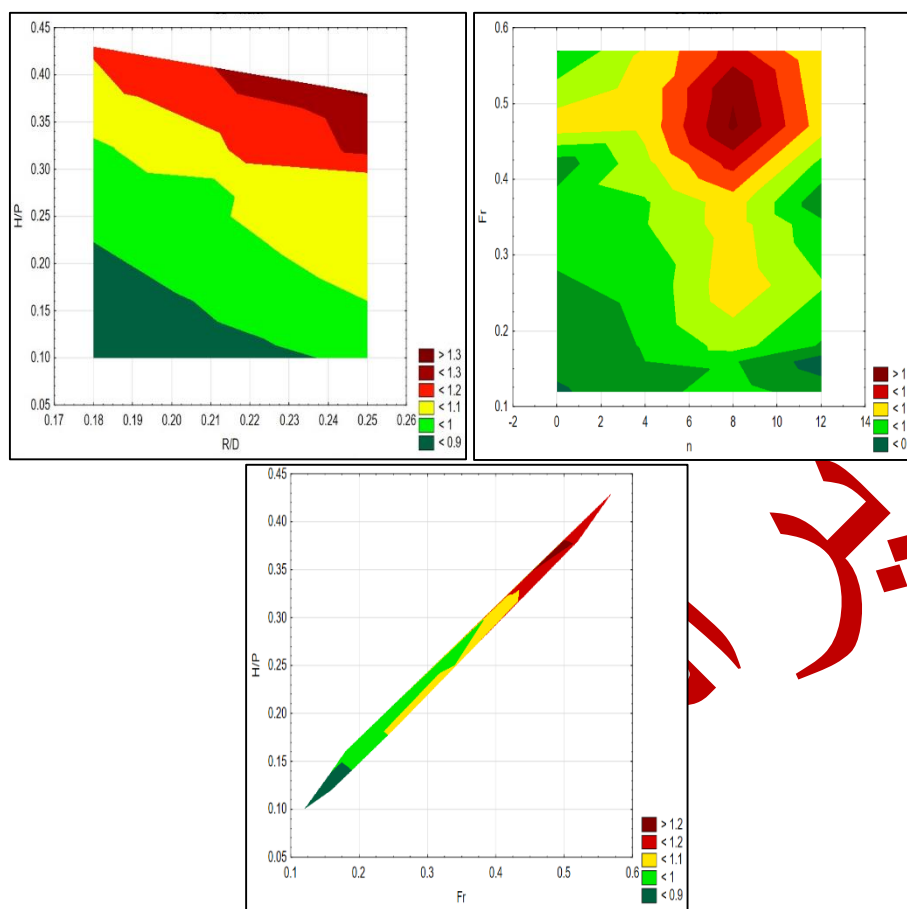
$(\frac{H}{p}, Fr)$ نشان می دهد.



شکل ۱۰. مقایسه بین نتایج مدل های بهینه SVM و GEP در گام آموزش (سمت راست) و آزمون (سمت چپ)



شکل ۱۱. بیان درختی از مدل GEP



شکل ۱۲. تغییرات ضریب دبی در برابر متغیرهای مستقل

نتیجه‌گیری

در این مقاله از داده‌های مدل فیزیکی سرریز نیلوفری-زیگزاکی با ورودی مربعی و دایروی برای تعیین ضریب دبی جریان تحت شرایط مختلف هندسی و هیدرولیکی با دو مدل *GEP* و *SVM* به منظور بررسی سناریوهای مختلف استفاده شد. داده‌های مورد استفاده در این پژوهش شامل n (تعداد زیگزاگ)، Fr (عدد فرود)، H/P (بار نسبی آبی) و R/D (شاخص شکل سرریز) بودند. در ابتدا با استفاده از ۸۰ داده آزمایشگاهی، فرایند آموزش و آزمون هر دو مدل انجام شد. سهم داده‌های آموزش و آزمون به ترتیب ۶۰ و ۲۰ داده بود. به کمک این داده‌ها، پارامترهای تنظیمی هر دو مدل برای دستیابی به بهینه‌ترین خروجی از طریق روش آزمون و خطا تعیین شدند. سپس خروجی هر یک از دو مدل با یکدیگر مقایسه شدند تا دقیق‌ترین مدل مشخص شود. سپس تحلیل حساسیت روی چهار متغیر مستقل انجام شد تا موثرترین پارامتر روی ضریب دبی که عملکرد هیدرولیکی سرریز را تحت تاثیر قرار می‌دهد مشخص شود. از بین تمام مدل‌های مورد بررسی در مورد *SVM*، تابع کرنل *RBF* به ازای $\gamma=0.1$ منجر به بهینه‌ترین خروجی شد به طوری که مقدار ($RMSE$, MAE , R^2) در دوره آموزش و آزمون برای این مدل به ترتیب (۰/۰۶۹۶، ۰/۹۲۶۲، ۰/۰۸۴۸) و (۰/۰۳۹۸، ۰/۰۳۴۶، ۰/۹۸۲۰) برای سرریز دایروی و (۰/۰۹۰۴، ۰/۰۷۳، ۰/۹۷۰۷) و (۰/۰۹۳۳۴، ۰/۰۶۷۶، ۰/۰۷۸۷) برای مقطع مربعی به دست آمدند. در خصوص مدل *GEP* نتایج بهتری به دست آمد به طوری که بهینه‌ترین خروجی به ازای مدل با سه ژن، اندازه هدا ۹ و تعداد کروموزم ۴۵ با شاخص‌های ارزیابی عملکرد (۰/۰۳۷۵، ۰/۰۴۵۱، ۰/۰۸۱۱) و (۰/۰۳۱۵، ۰/۰۳۹۶، ۰/۰۸۴۸) در گام‌های آموزش و آزمون در مقطع دایروی حاصل شدند. مقدار شاخص‌های ارزیابی عملکرد برای مقطع مربعی با ۵۵ کروموزم به ترتیب (۰/۰۴۹۴، ۰/۰۵۹۷، ۰/۰۷۴۱) و (۰/۰۵۹۴، ۰/۰۵۰۳، ۰/۰۹۵۹۱) در گام آموزش و آزمون تعیین شدند. با

توجه به مطالعه انجام شده، (۱) تحلیل مقایسه‌ای مدل‌های یادگیری ماشین بیشتر؛ (۲) بررسی تأثیر عوامل محیطی و اقلیمی؛ (۳) اثر عوامل خارجی مانند رسوب‌گذاری، تغییرات دما یا تأثیرات تغییرات اقلیمی بر عملکرد سرریز و ارزیابی ریسک آن‌ها؛ (۴) اعتبارسنجی با داده‌های میدانی به عنوان پیشنهاد برای پژوهش‌های آتی ارائه می‌شود.

منابع

منابع فارسی

- ابراهیم‌زاده، علی؛ ضرغامی، مهدی و نورانی، وحید (۱۳۹۸). مدیریت ریسک روگذری سد حاجیلر چای با شبیه‌سازی مونت کارلو و پویایی سیستم‌ها. *مجله مدیریت آب و آبیاری*، ۹(۲)، ۲۳۱-۲۵۰.
- اقبال‌زاده، سعید؛ قزل‌سوفلو، عباسعلی و علامتیان، جواد (۱۴۰۲). بازاندیشی طراحی بخش‌های مختلف سامانه سرریز بر اساس تحلیل ریسک چند سطحی (مطالعه موردی سد قزل داش). *نشریه آبیاری و زهکشی ایران*، ۱۷(۲)، ۱۹۳-۲۰۵.
- بهادری، خشایار و کریمایی طبرستانی، مجتبی (۱۳۹۹). تعیین ارتفاع و ریسک روگذری سدهای مخزنی بر مبنای تحلیل قابلیت اطمینان (مطالعه موردی: سد نمرود). *سد و نیروگاه برق آبی ایران*، ۷(۲۵)، ۱-۱۳.
- رضاپورطبری، محمود محمد و هاشم‌پور، معصومه (۱۳۹۷). توسعه مدل‌های هیبریدی PSO-DSO و GWO-DSO جهت باز طراحی ابعاد بهینه سرریز کنگره‌ای. *نشریه سد و نیروگاه برق آبی ایران*، ۵(۱۶)، ۴۸-۶۳.
- شرافتی، احد و و ذهبیون، باقر (۱۳۹۲). تحلیل ریسک روگذری سد با در نظر گرفتن عدم قطعیت‌های هیدرولوژیکی و هیدرولیکی. *هیدرولیک*، ۸(۱)، ۱۷-۱.
- فقیه، همایون؛ خلقی، مجید و کوچک‌زاده، صلاح (۱۳۸۷). ارزیابی و مقایسه تعدادی از روش‌های تحلیل کمی ریسک در برآورد سیلاب طراحی سرریز سدها (مطالعه موردی: سد پیشین). *نشریه علوم آب و خاک (علوم و فنون کشاورزی و منابع طبیعی)*، ۴(۱۲)، ۴۶۳-۴۷۴.
- فولادی‌پناه، مهدی و ماجدی‌اصل، مهدی (۱۴۰۱). کاربرد محاسبات نرم در افزایش دقت پیش‌بینی ضربه دبی جریان سرریزهای جانبی مستطیلی. *نشریه علمی-پژوهشی مهندسی آبیاری و آب ایران*، ۱۲(۴۸)، ۲۱۳-۲۳۳.
- فولادی‌پناه، مهدی؛ ماجدی‌اصل، م هدی و حق‌گویی، آیدا. (۱۳۹۹). کاربرد الگوریتم‌های هوشمند برای مدل‌سازی رابطه دبی-اشل در شرایط استغراق سرریزهای کنگره‌ای و خطی. *نشریه هیدرولیک*، ۱۵(۲)، ۱۴۹-۱۶۴.
- فیضی، احسان؛ نقوی، محمد و فخرایی، حسین (۱۳۹۹). ارزیابی ریسک سدهای بتنی برآبی با روش تلفیقی FEMA و RAMCAP با رویکرد پدافند غیرعامل (مطالعه موردی: سد بتنی لیرو). *نشریه علمی پدافند غیرعامل*، ۱۱(۲)، ۸۳-۹۴.
- کاراموز، محمد؛ درودی، سیامک و مریدی، علی (۱۳۹۵). طراحی بهینه برای ابعاد سیستم انحراف آب در سدها با بررسی و تألیف عدم قطعیت هیدرولیکی و ریسک هیدرولوژیکی. *هیدرولیک*، ۱۱(۱)، ۲۱-۳۴.
- مغربی، محسن؛ قزل‌سوفلو، عباسعلی و علیمیرزایی، حسین (۱۳۹۶). ارزیابی ریسک در اجزای سازه‌ی سرریز (مطالعه موردی سد چندیر). *آب و توسعه پایدار*، ۴(۲)، ۴۱-۴۸.

منابع انگلیسی

- Aouissi, H.A., Petrisor, A.I., Ababsa, M., Bo-stenaru-Dan, M., Tourki, M., & Bouslama, Z. (2021). Influence of Land Use on Avian Diversity in North African Urban Environments. *Land*, 10(4), 434. <https://doi.org/10.3390/land10040434>
- Bahadori, K.H., & Karimaei Tabarestani, M. (2020). Determination of the Height and Overtopping failure of Reservoir Dams by Using Reliability Analysis (Case Study: Namrood Dam). *Journal of Iranian Dam and Hydropower*, 25(7), 1-13. (In Persian)
- Borowski, P.F. (2020). New technologies and innovative solutions in the development strategies of energy enterprises. *HighTech and innovation Journal*, 1(2), 39-58. DOI: 10.28991/HIJ-2020-01-02-01
- Ebrahimzadeh, A., Zarghami, M., & Nourani, V. (2020). Overtopping risk management by system dynamics and Monte-Carlo simulations, Hajilarchay Dam of Iran. *Water and Irrigation Management*, 9(2), 231-250. DOI: 10.22059/jwim.2019.290802.719 (In Persian)

- Eghbalizadeh, S., Ghezelsofloo, A.A., & Alamatian, J. (2023). Rethinking the design of different parts of the free overflow system based on multi-level risk analysis (Case Study Ghezel dash). *Iranian Journal of Irrigation & Drainage*, 17(2), 193-205. DOI: 20.1001.1.20087942.1402.17.2.1.7 (In Persian)
- Faghih, H., Kholghi, M., & Kochekezadeh, S. (2009). Evaluating and Comparing Some of the Quantitative Risk Analysis Methods to Estimate Design Flood of Dam Spillway (Case Study: Pishin Dam Spillway). *Journal of Crop Production and Processing*, 12(46), 463-474. DOI: 20.1001.1.22518517.1387.12.46.5.9 (In Persian)
- Ferreira, C. (2001). Algorithm for solving gene expression programming: a new adaptive problem. *Complex Systems*, 13(2), 87-129. DOI: 10.4236/ajor.2018.82008
- Feyzi, E., Naghavi, M., & Fakhraei, H. (2020). Risk Assessment of Hydroelectric Concrete Dams Using Combined FEMA and RAMCAP Method with Passive Defense Approach, Case Study: Leero Concrete Dam. *Passive Defense*, 11(2), 83-94. DOI:20.1001.1.20086849.1399.11.2.8.9 (In Persian)
- Frizell, K.W., Renna, F.M., & Matos, J. (2013). Cavitation potential of flow on stepped spillways. *Journal of Hydraulic Engineering*, 139(6), 630-636. [https://doi.org/10.1061/\(ASCE\)HY.1943-7900.0000715](https://doi.org/10.1061/(ASCE)HY.1943-7900.0000715)
- Fuladipناه, M., Majedi Asl, M., & Haghgooyi, A. (2020). Application of intelligent algorithm to model head-discharge relationship for submerged labyrinth and linear weirs. *Journal of Hydraulics*, 15(2), 149-164. DOI:10.30482/jhyd.2020.232388.1461 (In Persian)
- Fuladipناه, M., & Majedi-Asl, M. (2022). Soft Computing Application to Amplify Discharge Coefficient Prediction in Side Rectangular Weirs. *Irrigation and Water Engineering*, 12(4), 213-233. DOI:10.22125/iwe.2022.150692 (In Persian)
- Gaagai, A., Aouissi, H.A., Krauklis, A.E., Burlakovs, J., Athamena, A., Zekker, I., Boudoukha, A., Benaabidate, A., & Chenchouni, H. (2022). Modeling and risk analysis of dam-break flooding in a semi-arid Montane watershed: a case study of the Yabous Dam, Northeastern Algeria. *Water*, 14(5), 767-797. <https://doi.org/10.3390/w14050767>
- Karamouz, M., Doroudi, S., & Moridi, A. (2016). An Optimal Design for Dimensions of Water Diversion System in Dams using and Analyzing Hydraulic Uncertainties and Hydrologic Risk. *Journal of Hydraulics*, 11(1), 21-34. DOI: 10.30482/jhyd.2016.41484. (In Persian)
- Lerma, N., Paredes-Arquiola, J., Andreu, J., Solera, A., & Sechi, G.M. (2015). Assessment of evolutionary algorithms for optimal operating rules design in real water resource systems. *Environmental Modelling and Software*, 69, 425-436. <https://doi.org/10.1016/j.envsoft.2014.09.024>
- Lucas, J., Hager, W.H., & Boes, R.M. (2013). Deflector effect on chute flow. *Journal of Hydraulic Engineering*, 139(4), 444-449. [https://doi.org/10.1061/\(ASCE\)HY.1943-7900.0000652](https://doi.org/10.1061/(ASCE)HY.1943-7900.0000652)
- Maghrebi, M., Ghezelsofloo, A., & Alimirzaei, H. (2018). Risk Assessment for Spillway Overflow Structure Components (Case Study: Chandir Dam). *Journal of Water and Sustainable Development*, 4(2), 41-48. DOI:10.22067/jwsd.v4i2.59241 (In Persian)
- Rezapour Tabari, M.M., & Hashempour, M. (2018). Development of GWO-DSO and PSO-DSO Hybrid Models to Redesign the Optimal Dimensions of Labyrinth Spillway. *Journal of Iranian Dam and Hydropower*, 5(16), 48-63. DOI:10.1007/s00500-018-3292-9 (In Persian)
- Pfister, M., Lucas, J., & Hager, W.H. (2011). Chute aerators: preaerated approach flow. *Journal of Hydraulic Engineering*, 137(11), 1452-1461. [https://doi.org/10.1061/\(ASCE\)HY.1943-7900.0000417](https://doi.org/10.1061/(ASCE)HY.1943-7900.0000417)
- Sedighizadeh, S. (2011). A new model for economic optimization of water diversion system during dam construction using PSO algorithm. *J. World Academy of Science, Engineering and Technology*, 5, 2011-2020.
- Sharafati, A., & Zahabiyou, B. (2013). Analysis of Dam Overtopping by Considering Hydraulic and Hydrological Uncertainties. *Journal of Hydraulics*, 8(1), 1-17. DOI: 10.30482/jhyd.2014.7472 (In Persian)
- Vapnik, V. (1995). *The Nature of Statistical Learning Theory*. Springer-Verlag. New York.

Extended Abstract

Introduction

Project risk encompasses unforeseen events affecting time, cost, and quality. Effective management involves analysis and mitigation, crucial in engineering, particularly dam design. Karamoz et al. (2015) examined water diversion system design challenges. Maghrebi et al. (2016) assessed spillway risks for Chandir Dam. Bahadori and Karimai-Tabarestani (2019) analyzed reservoir dam height and crossing risks. Faizi et al. (2019) used FEMA and RAMCAP methods for Liro dam risk evaluation. Iqbalizadeh et al. (2023) advocated multi-level risk analysis for spillway redesign, while Rezapour and Hashempour (2017) proposed hybrid models for optimizing spillway dimensions. Lakos et al. (2020) and Frizel et al. (2020) highlighted the need for safe, economical spillways due to large dam construction and higher safety standards.

The literature review confirms extensive research in risk assessment, focusing on hydraulic and hydrological scenarios. However, the forthcoming study uniquely applies MLMs, including SVM and GEP to evaluate the Cd of the glory-labyrinth spillway. This is achieved by introducing labyrinth configurations at the inlet inlet—a novel approach not previously explored in existing studies.

Materials and Methods

In this study, a comprehensive approach was employed to simulate the Cd of the glory-labyrinth spillway utilizing 80 laboratory data sets collected from two distinct inlet sections featuring square and circular labyrinth configurations. These configurations varied in the number of labyrinths, specifically four, eight, and twelve. To accurately model the Cd, two advanced MLMs were implemented: SVM and GEP. The independent variables considered in the simulations included the number of labyrinth (n), Froude number (Fr), relative water load (H/P), and the weir shape index (R/D). These variables were chosen due to their critical influence on the hydraulic behavior of the weir. To assess the accuracy and reliability of the models, performance evaluation indices, namely Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2), were employed. These indices provided a quantitative measure of the models' predictive capabilities, ensuring that the simulated results closely align with the observed data.

Results

In the evaluation of various SVM models, the RBF kernel function with γ set to 0.1 yielded the most optimal results. The model's performance metrics (RMSE, MAE, R^2) during the training and testing phases were (0.9262, 0.0696, 0.0848) and (0.9820, 0.0346, 0.0398) for the circular spillway, and (0.9707, 0.073, 0.0904) and (0.9334, 0.0676, 0.0787) for the square section. Superior results were obtained using the GEP model, particularly with three genes, a head size of 9, and 45 chromosomes. For the circular spillway, the GEP model achieved indices of (0.9778, 0.0375, 0.0451) and (0.9811, 0.0315, 0.0396) during training and testing, respectively. In the square section, the model with 55 chromosomes showed performance values of (0.9741, 0.0494, 0.0597) and (0.9591, 0.0503, 0.0594) in the training and testing phases, respectively.

Discussion and Conclusion

The evaluation of various SVM models identified the RBF kernel function with a specific γ value as yielding the most optimal results. The model's performance was assessed for both circular and square spillways, showing strong metrics in both training and testing phases. Additionally, the GEP model, particularly with specific genetic configurations, demonstrated superior performance across different spillway geometries, in

Keywords: Diverted Flow; Computational Fluid Dynamics; Data-Driven model; Performance Assessment.

Author Contributions

All authors contributed equally to the conceptualization of the article and writing of the original and subsequent drafts.

Data Availability Statement

Data available on request from the authors.

Acknowledgements

The authors would like to thank all participants of the present study.

Ethical considerations

The authors avoided data fabrication, falsification, plagiarism, and misconduct.

Conflict of interest

The author declares no conflict of interest.