

## کاربرد هوش مصنوعی در پیش‌بینی ضریب دبی در سرریزهای کنگره‌ای با پلان هارمونیک

امل سواعدی<sup>۱</sup>، محمدرضا زایری<sup>۲\*</sup>، مهدی قمشی<sup>۳</sup>، مهدی دریائی<sup>۴</sup>

<sup>۱</sup> دانشجوی کارشناسی ارشد هیدروانفورماتیک، گروه سازه‌های آبی، دانشکده مهندسی آب و محیط‌زیست، دانشگاه شهید چمران اهواز، اهواز، ایران.  
<sup>۲\*</sup> به ترتیب استادیار، استاد و دانشیار گروه سازه‌های آبی، دانشکده مهندسی آب و محیط‌زیست، دانشگاه شهید چمران اهواز، اهواز، ایران.

تاریخ دریافت: ۱۴۰۳/۰۹/۲۹ تاریخ پذیرش: ۱۴۰۳/۱۰/۱۵

### چکیده

با رشد جمعیت، تقاضا برای ذخیره‌سازی آب در مخازن افزایش یافته است و نیاز به جایگزینی یا افزایش ظرفیت دبی بسیاری از سرریزها از طریق بهینه‌سازی هندسی و طرح افقی آنها به وجود آمده است. یکی از راهکارهای کاربردی برای افزایش راندمان سرریزها، اصلاح هندسه پلان و افزایش طول سرریز در عرضی ثابت است. این امر باعث افزایش ضریب دبی (Ca) سرریز می‌شود. در این مطالعه، با جمع‌آوری داده‌های آزمایشگاهی دو منبع مختلف، در مجموع از ۲۳۳ سری داده شامل مؤلفه‌های هندسی و هیدرولیکی استفاده شد. سه مدل یادگیری ماشین، شامل رگرسیون بردار پشتیبان (SVR) و مدل ناپارامتریک جنگل تصادفی (RF) و رگرسیون اسپیرالین تطبیقی چندگانه (MARS)، با پنج سناریوی ورودی مختلف برای پیش‌بینی دقیق ضریب دبی سرریز کنگره‌ای با پلان نیم‌دایره هارمونیک توسعه داده شد. نتایج تحقیق نشان داد که متغیرهای ورودی شامل، نسبت شعاع به ارتفاع سرریز (R/P)، نسبت عرض سرریز به ارتفاع سرریز (B/P)، تعداد سیکل (N) و نسبت هد هیدرولیکی به ارتفاع سرریز (H<sub>T</sub>/P)، در مدل MARS با مقادیر  $R^2=0/94$  و  $RMSE=0/008$  و مدل SVR با مقادیر  $R^2=0/93$  و  $RMSE=0/009$ ، نتایج بهتری را در تخمین (Ca) ارائه می‌کنند. مشاهده گردید در مقادیر (H<sub>T</sub>/P) پایین‌تر از ۰/۰۵، میزان ضریب دبی عبوری سرریز با ۴ سیکل نیم‌دایره (N=۴)، بیشتر از سایز سرریزها بوده است. همچنین، با استفاده از دو روش مختلف آنالیز حساسیت مشاهده گردید پارامتر عمق نسبی (H<sub>T</sub>/P) مؤثرترین مؤلفه در پیش‌بینی ضریب دبی بوده است.

واژه‌های کلیدی: آنالیز حساسیت SHAP، یادگیری ماشین، ضریب دبی، سرریزهای کنگره‌ای

### مقدمه

راه‌حل‌های ممکن برای کاهش این مشکل، استفاده از سرریز غیرخطی یا همان کنگره‌ای<sup>۱</sup> است. این نوع سرریزها نه تنها ظرفیت دبی جریان را افزایش می‌دهند، بلکه می‌توانند از افزایش عمق ناشی از سیلاب نیز جلوگیری کنند. این سرریزها همچنین به عنوان سازه‌های مؤثر در افت انرژی و هوادهی جریان و سازه‌های استهلاک انرژی نیز به کار می‌روند (Ghaderi et al., 2024; Hussein & Jalil, 2024; Ikinciogullari, 2024)

سدها که به عنوان موانعی در مسیر رودخانه قرار می‌گیرند، از اجزای اصلی مانند بدنه، سرریز و زهکش ساخته می‌شوند. یکی از عواملی که به شکست سدها کمک می‌کند، ظرفیت ناکافی به‌منظور تخلیه سیلاب از مخزن سد است. انتخاب نوع سرریز معمولاً مسئله‌ای بهینه‌سازی است که هدف آن حداکثرکردن ظرفیت تخلیه دبی از سرریز با کمترین هزینه ساخت است. طول و شکل تاج سرریز دو عامل اصلی هستند که ظرفیت دبی آن را تعیین می‌کنند. یکی از

<sup>۱</sup> Labyrinth weirs



عددی با استفاده از نرم‌افزار FLOW-3D مقایسه کردند. مدل‌های عددی آن‌ها به دقتی تا ۹۶ درصد، در مقایسه با داده‌های آزمایشگاهی، دست یافتند که نشان‌دهنده کارایی مدل‌سازی عددی در پیش‌بینی دبی سرریز است. این محققان همچنین مشاهده کردند که عملکرد ضریب دبی، در مقایسه با به نسبت هد کل به ارتفاع سرریز ( $H_T/P$ )، برای سرریزهای نیم‌دایره‌ای مشابه سرریزهای دوزنقه‌ای با زوایای دیواره جانبی بیشتر از ۳۵ درجه است. سنگ سفیدی و همکاران (Sangsefidi *et al.*, 2018) در آزمایشگاه روی سرریز کنگره‌ای دایره‌ای شکل دنداندار مطالعه کردند و به این نتیجه رسیدند که ظرفیت دبی این نوع از سرریزها، در مقایسه با سرریز خطی،  $4/5$  برابر افزایش یافته است. دووتا و همکاران (Dutta *et al.*, 2020) عملکرد پیش‌بینی دبی در شکل‌های سرریز (سرریز و نیم‌دایره‌ای) را با استفاده از رگرسیون خطی چندگانه (MLR)، ماشین‌های بردار پشتیبان (SVR)، و شبکه‌های عصبی مصنوعی (ANN) ارزیابی کردند و بر اساس شاخص‌های آماری مانند ضریب تعیین ( $R^2$ )، میانگین درصد خطا (APE) و ضریب کارایی (CE)، نتیجه گرفتند که مدل SVR عملکرد بهتری نسبت به دیگر روش‌ها دارد که نشان‌دهنده توانایی برتر آن در درک روابط پیچیده بین پارامترهای دبی جریان دبی برای این هندسه‌های سرریز است. شفیع و همکاران (Shafiei *et al.*, 2020b) ضریب دبی تاج سرریزها را با استفاده از یک روش مبتنی بر هوش مصنوعی (AI) تخمین زدند. نتایج پژوهش‌های آن‌ها نشان داد که با وجود محدودیت‌هایی مانند نبود معادله‌ای قطعی، مدل ORELM عملکرد قابل قبولی در شبیه‌سازی ضریب دبی داشته است. شفیع و همکاران (Shafiei *et al.*, 2020a) از سیستم استنتاج عصبی-فازی تطبیقی (ANFIS) به همراه الگوریتم کرم شب تاب (FFA) برای تخمین ضریب دبی تاج سرریزهای مثلثی استفاده کردند و نتیجه گرفتند که مدل ANFIS-FFA، نسبت به مدل استاندارد ANFIS، در پیش‌بینی ضریب دبی تاج

سرریزهای مثلثی عملکرد بهتری داشته است. بنکداری و همکاران (Bonakdari *et al.*, 2020) از تکنیک برنامه‌نویسی بیانی ژن (GEP) برای تخمین ضریب دبی تاج سرریزهای مثلثی استفاده کردند و نشان دادند تکنیک GEP در پیش‌بینی ضریب دبی تاج سرریزهای مثلثی نتایج رضایت‌بخشی ارائه کرده است. امامی و همکاران (Emami *et al.*, 2021) یک الگوریتم بهبود یافته تکامل تفاضلی خودتطبیقی (ISADE) را همراه با رگرسیون بردار پشتیبان (SVR\_ISADE) برای تخمین ضریب دبی تاج سرریزهای W شکل توسعه دادند. نتایج مطالعات آن‌ها نشان داد که روش SVR\_ISADE در پیش‌بینی ضریب دبی تاج سرریزهای W شکل عملکرد برتری دارد به طوری که مقدار  $R^2$  برابر  $0/982$ ، MAPE برابر  $0/612$ ، RMSE برابر  $0/006$  و  $\delta$  برابر  $0/843$  بوده است. صفر رضوی‌زاده و همکاران (Safarrazavi Zadeh *et al.*, 2021) مطالعه‌ای آزمایشگاهی برای بررسی ضریب دبی جریان تحت تأثیر نسبت هد هیدرولیکی بالادست ( $H_T/P$ )، نسبت طول سرریز ( $L/P$ )، نسبت شعاع قوس ( $R/P$ ) و تعداد سیکل‌های سرریز ( $N$ ) برای سرریزهای کنگره‌ای نیم‌دایره‌ای و سینوسی کردند. نتایج مطالعات این پژوهشگران نشان داد که برای نسبت‌های ( $H_T/P < 0/35$ )، سرریزهای کنگره‌ای نیم‌دایره‌ای و سینوسی روندی افزایشی در ضریب دبی داشته‌اند که سپس کاهش می‌یابد. این محققان همچنین دریافتند که دبی جریان عبوری از سرریزهای کنگره‌ای نیم‌دایره‌ای و سینوسی ۳۳ درصد بیشتر از دبی جریان عبوری از سرریزهای خطی است. سلیمانی و همکاران (Suleimany *et al.*, 2022) از رگرسیون خطی چندگانه (MLR) و رگرسیون قدرتی برای توسعه معادله‌ای تجربی برای دبی سرریز نیم‌دایره‌ای استفاده کردند. مدل‌های این محققان به ضریب تعیین ( $R^2$ ) بالایی به مقادیر  $0/8498$  و  $0/8584$  رسیدند که نشان‌دهنده تطابق خوب با داده‌های تجربی است. حقیبن و همکاران (Haghibin *et al.*, 2022) یک رویکرد مبتنی بر داده برای تخمین ضریب دبی

## مواد و روش‌ها

### تحلیل ابعادی

ضریب دبی ( $C_d$ ) سرریزهای کنگره‌ای نیم‌دایره‌ای را می‌توان با استفاده از معادله ۱ تعیین کرد. در این مغداله،  $Q$  نشان‌دهنده دبی،  $C_d$  ضریب دبی،  $L$  طول سرریز،  $g$  شتاب ناشی از گرانش، و  $H_t$  هد کل هیدرولیکی است:

$$Q = \frac{2}{3} C_d L \sqrt{2gH_t}^{1.5} \quad (1)$$

مطابق شکل (۲)، مؤلفه‌های مؤثر بر ضریب دبی سرریزهای کنگره‌ای نیم‌دایره‌ای را می‌توان به شرح زیر در نظر گرفت:

$$Cd = F_1(N, R, L, B, H_T, P, g, \rho, \mu, \delta) \quad (2)$$

که در آن،  $(N)$  تعداد سیکل کنگره ای،  $(R)$  شعاع سیکل‌های نیم‌دایره،  $(B)$  عرض کانال،  $(P)$  ارتفاع سرریز،  $(\rho)$  چگالی سیال،  $(\mu)$  ویسکوزیته دینامیکی، و  $(\delta)$  کشش سطحی است. به دلیل غیرقابل تراکم بودن جریان در سرریزها، اثر تغییرات جرم مخصوص ( $\rho$ ) قابل صرف نظر است. اثر گرانشی و کشش سطحی نیز اگر جریان به صورت آشفته باشد و ارتفاع جریان روی سرریز بر اساس توصیه رو و شاکلا (Rao & Shukla, 1971)، بیشتر از سه سانتی‌متر باشد، می‌تواند نادیده گرفته شود. بنابراین، با استفاده از تحلیل ابعادی، معادله ۲ می‌تواند به شکل زیر بازنویسی شود:

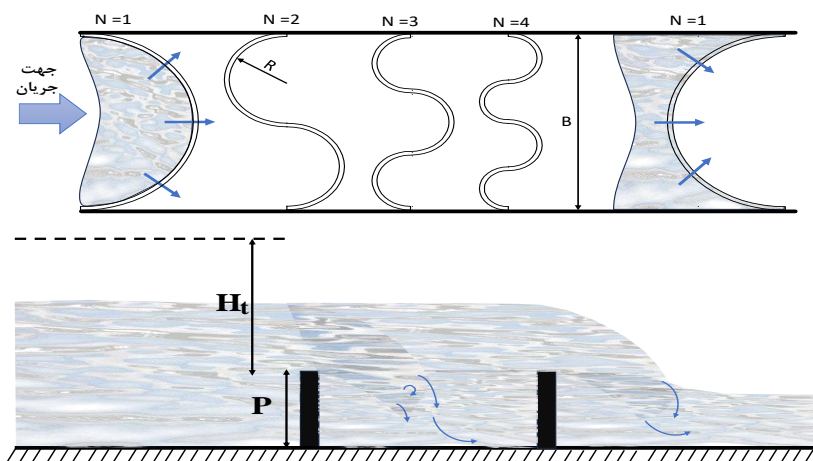
$$Cd = F_2\left(N, \frac{R}{P}, \frac{L}{P}, \frac{B}{P}, \frac{H_T}{P}\right) \quad (3)$$

در این مطالعه، به منظور توسعه مدل‌های محاسبات نرم از مجموعه نتایج آزمایشگاهی ارهام نامازی و مظفری (Arham Namazi & Mozaffari, 2023) و یلدیز و همکاران (Yıldız et al., 2024) استفاده شده است.

در سرریزهای پله‌ای پیشنهاد دادند. مدل بهینه‌سازی شده آن‌ها شاخص عملکرد را به میزان ۸۶/۱۳ درصد بهبود بخشید که نشان‌دهنده پتانسیل روش‌های مبتنی بر داده برای افزایش دقت پیش‌بینی دبی است. لی و همکاران (Li et al., 2024) از ماشین‌های بردار پشتیبان (SVR) به همراه سه الگوریتم بهینه‌سازی برای ایجاد یک مدل پیش‌بینی ضریب دبی ( $C_d$ ) در سرریزهای جانبی نیم‌دایره‌ای (SCSWs) استفاده و نتیجه‌گیری کردند که روش‌های SVR و الگوریتم ژنتیک (GASVR) دقت بالایی در پیش‌بینی و توانایی تعمیم دارند. میانگین خطا و حداکثر خطا به ترتیب ۰/۰۸ و ۰/۴۷ درصد بود که نسبت به مدل تجربی متداول به ترتیب ۷۲/۹۵ و ۸۶/۶۰ درصد کاهش یافته است.

با توجه به پارامترهای فیزیکی پیچیده سازه سرریز کنگره‌ای با طرح هارمونیک، به منظور پیش‌بینی ضریب دبی، انتخاب مدل‌های هوش مصنوعی دارای دقت و پایداری بالا ضروری است. تحقیقات نشان می‌دهد تاکنون مدل پیش‌بینی دقیق برای ضریب دبی ( $C_d$ ) این نوع سازه توسعه نیافته است. از این رو، این تحقیق مدل پیش‌بینی ضریب دبی ( $C_d$ ) برای سرریز با طرح هارمونیک را با روش‌های مختلف شامل رگرسیون ماشین بردار پشتیبان (SVR) و مدل ناپارامتریک جنگل تصادفی (RF) ارائه می‌دهد و افزون بر آن روشی جدید با استفاده از رگرسیون اسپلاین تطبیقی چندگانه (MARS) برای محاسبه دقیق ضریب دبی این سازه معرفی می‌کند. با جمع‌آوری مجموعه نمونه داده‌های دو مطالعه آزمایشگاهی، تأثیر پارامترهای بدون بعد ورودی به مدل بر ضریب دبی مورد تحلیل قرار گرفت. مدل‌های توسعه یافته در این مطالعه، با استفاده از زبان برنامه‌نویسی پایتون صورت گرفته است.

## کاربرد هوش مصنوعی در پیش‌بینی ضریب دبی در سرریزهای کنگره‌ای با پلان هارمونیک



شکل ۲- پارامترهای موثر بر ضریب دبی سرریز کنگره‌ای طرح هارمونیک (الف) نمای بالا (ب) نمای جانبی سرریز

Fig. 2-Importance Parameters on the discharge coefficient of the harmonic plan weirs (a) top view (b) side view

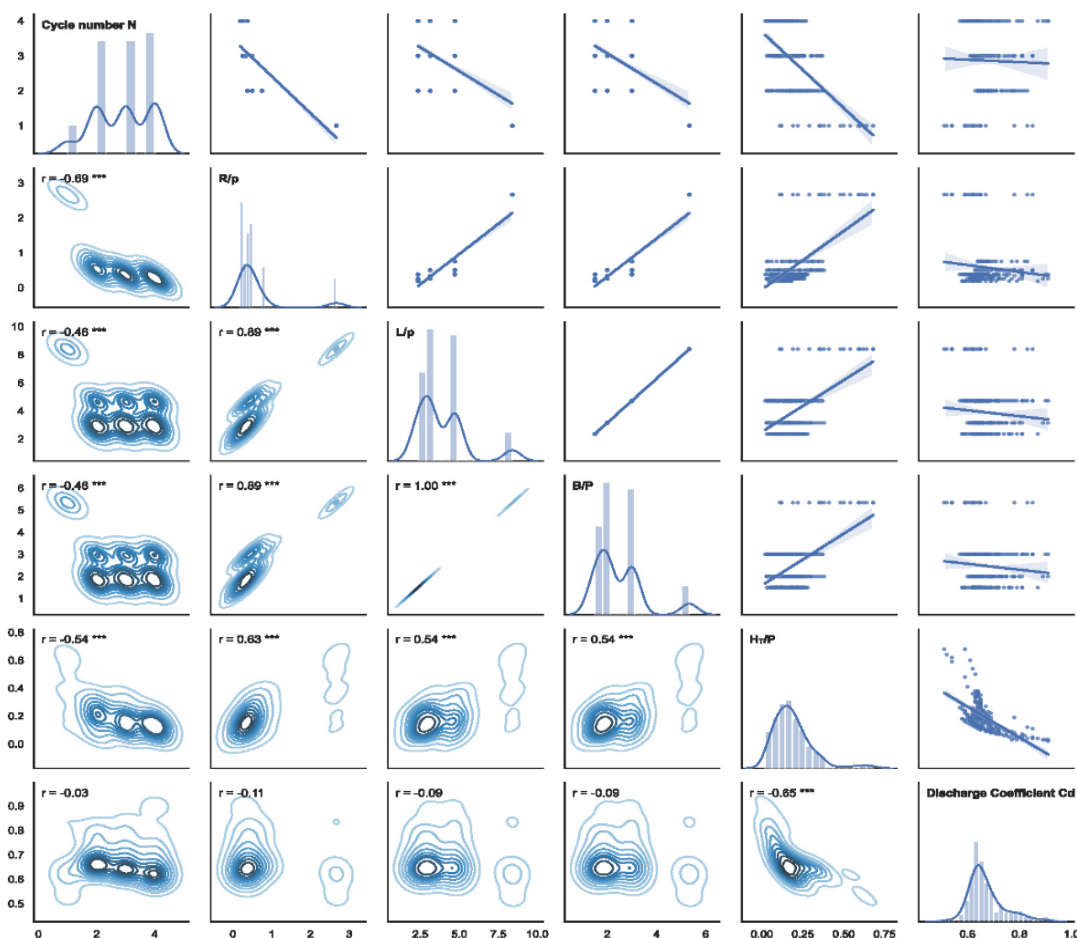
نرم جمع‌آوری گردید. از این تعداد، نتایج ۱۷۵ نمونه (۷۵ درصد) برای آموزش و ۵۸ نمونه (۲۵ درصد) برای آزمایش مدل‌های توسعه‌یافته در نظر گرفته شده است. در جدول (۱)، شاخص‌های آماری این دو بخش آورده شده است. شکل ۳ توزیع هیستوگرام و ضریب همبستگی بین متغیرهای مورد استفاده در پیش‌بینی  $C_d$  را نشان می‌دهد. بیشترین مقدار ضریب همبستگی بین متغیر  $C_d$  با متغیر عمق نسبی با مقدار ۰/۶۵- نشان‌دهنده معادله معکوس بین این دو متغیر است. به‌طورکلی، بین متغیرهای ورودی و ضریب دبی  $C_d$  همبستگی خطی ضعیفی وجود دارد، بنابراین پیش‌بینی دقیق بر اساس روش‌های کلاسیک رگرسیون دشوار است.

به منظور ارزیابی دقت معادلات پیشنهادی، مدل‌های محاسبات نرم در برآورد ضریب دبی از سرریزهای کنگره‌ای دایره‌ای که به‌صورت هارمونیک در کانال روباز قرار داده شده‌اند، در مطالعات بلدیز و همکاران (Yildiz *et al.*, 2024) برای سه ارتفاع مختلف سرریز ( $P=20$  سانتی‌متر،  $P=30$  سانتی‌متر و  $P=40$  سانتی‌متر) و سه تعداد مختلف سیکل ( $N=2$ ،  $N=3$  و  $N=4$ )، به تعداد ۲۱۵ آزمایش و در مطالعات ارهام نمازی و مظفری و Mozaffari (Arham Namazi & Mozaffari, 2023) برای ارتفاع ثابت سرریز  $P=15$  سانتی‌متر که به‌صورت تک سیکل ( $N=1$ ) به تعداد ۱۸ آزمایش اجرا شده است که در مجموع ۲۳۳ آزمایش برای مدل‌سازی مبتنی بر محاسبات

جدول ۱- آماره‌های توصیفی داده‌های تخصیص داده‌شده در مرحله آموزش و اعتبارسنجی مدل‌های محاسبات نرم

Table 1- Statistical characteristics of data assigned to Training and Test soft computing models

H/P	B/P	L/p	R/p	N	$C_d$	شاخص	مرحله
۰/۰۲۰	۱/۵۰۰	۲/۳۵۶	۰/۱۸۷	۱/۰۰۰	۰/۵۱۰	کمینه	آموزش
۰/۶۸۰	۵/۳۳۳	۸/۴۰۰	۲/۶۶۷	۴/۰۰۰	۰/۹۱۰	بیشینه	
۰/۱۸۷	۲/۴۱۴	۳/۷۹۴	۰/۵۴۶	۲/۹۰۳	۰/۶۷۲	میانگین	
۰/۱۲۰	۰/۹۸۳	۱/۵۴۸	۰/۵۹۶	۰/۹۵۱	۰/۰۷۱	انحراف معیار	
۰/۰۳۰	۱/۵۰۰	۲/۳۵۶	۰/۱۸۷	۱/۰۰۰	۰/۵۴۰	کمینه	اعتبارسنجی
۰/۶۲۰	۵/۳۳۳	۸/۴۰۰	۲/۶۶۷	۴/۰۰۰	۰/۸۹۰	بیشینه	
۰/۱۹۲	۲/۶۵۵	۴/۱۷۳	۰/۶۶۵	۲/۷۵۹	۰/۶۷۲	میانگین	
۰/۱۱۳	۱/۰۹۸	۱/۷۳۱	۰/۷۰۸	۰/۹۷۹	۰/۰۶۹	انحراف معیار	



شکل ۳- نمودار هیستوگرام و ماتریس همبستگی بین همه متغیرهای ورودی و ضریب دبی

Fig. 3- Correlation matrix diagram of input and output variables

### مدل ماشین بردار پشتیبان

بنابراین، هدف اصلی یافتن فرم تابع  $f$  است که بتواند به درستی موارد جدیدی را پیش‌بینی کند که SVR تاکنون ندیده است. در روش  $\epsilon$ -SVR، از توابع هسته‌ای (کرنل) برای جداسازی صفحات و بهینه‌سازی فاصله‌ها میان داده‌های مشاهده‌شده و تخمینی استفاده می‌شود. اما در مدل V-SVR از برنامه‌ریزی غیرخطی برای کاهش خطای پیش‌بینی بهره گرفته می‌شود (Behdarvandi *et al.*, 2024). در این مطالعه، به دلیل کاربرد وسیع مدل  $\epsilon$ -SVR در مسائل رگرسیونی، از آن استفاده شده است. تابع خطا برای این مدل به شکل زیر تعریف می‌شود:

$$\frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^I (\xi_i^- + \xi_i^+) \quad (5)$$

ماشین بردار پشتیبان (SVM) سیستمی کارآمد برای یادگیری است که بر مبنای نظریه بهینه‌سازی مقید طراحی شده و از اصل کمینه‌سازی خطای ساختاری پیروی می‌کند تا به یک پاسخ بهینه کلی دست یابد. در مدل رگرسیون SVR، لازم است وابستگی متغیر هدف ( $y$ ) به مجموعه‌ای از متغیرهای مستقل ( $x$ ) برآورد شود. فرض می‌شود که مانند دیگر مدل‌های رگرسیون، رابطه میان متغیر هدف و متغیرهای مستقل از طریق یک تابع مشخص  $f$  و یک مقدار اضافی نویز، مطابق معادله ۴ تعریف می‌شود.

$$y = f(x) + \text{noise} \quad (4)$$



SVR تلاش می‌کند مرزی را پیدا کند که حاشیه بین کلاس‌ها را بیشینه کند.

$$y = f(x) = \text{sign} \left[ \sum_{i=1}^N y_i a_i K(x, x_i) + b \right] \quad (9)$$

پارامتر (K) تابع کرنل است که ضرب‌های داخلی ایجاد کرده است و امکان ساخت ماشین‌هایی با سطوح غیرخطی مختلف را در فضای داده‌ها فراهم می‌کند. برای مدل رگرسیون ماشین بردار پشتیبان، از کرنل‌های گوناگونی مانند خطی، درجه دوم، گوسی و چندجمله‌ای استفاده می‌شود. معمولاً توابع کرنل گوسی شعاعی (RBF) یا خطی در پیش‌بینی عملکرد بهتری دارند (Zayeri, 2023). معادله تابع کرنل RBF به صورت معادله ۱۰ ارائه می‌شود.

$$K(x, x_i) = \exp \left( -\frac{\|x - x_i\|^2}{2\sigma^2} \right) \quad (10)$$

### مدل جنگل تصادفی

مدل جنگل تصادفی (RF) با سرعت بالا طبقه‌بندی را برای بسیاری از مجموعه‌داده‌ها اجرا می‌کند. برخلاف مدل‌های دیگر، مانند رگرسیون چندمتغیره، این مدل از صدها یا هزاران درخت استفاده می‌کند تا از اطلاعات موجود در داده‌ها به طور گسترده‌تری بهره‌برداری و بهترین استنباط را از متغیرهای ورودی فراهم کند. این روش، یکی از مدل‌های ناپارامتری رایج است که از تکنیک بگینگ<sup>۵</sup> استفاده می‌کند، ترکیبی از چندین درخت تصمیم که برای ساخت آن، نمونه‌های بوت‌استرپ<sup>۶</sup> متعدد از داده‌ها انتخاب می‌شوند و در هر درخت، تعدادی از متغیرهای ورودی به طور تصادفی در نظر گرفته می‌شود. در نتیجه، برخی از داده‌ها ممکن است چندین بار در فرآیند آموزش به کار روند، درحالی‌که برخی دیگر که در مدل‌سازی مؤثر نیستند، هرگز استفاده نمی‌شوند.

لازم است تابع خطای مذکور با در نظر گرفتن محدودیت‌های زیر به حداقل برسد:

$$y_i - ((\omega, x_i) + b) \leq \varepsilon + \xi_i, \xi_i \geq 0 \quad (6)$$

$$((\omega, x_i) + b) - b \leq \varepsilon + \xi_i^*, \xi_i^* \geq 0 \quad (7)$$

در این معادلات، C یک مقدار از پیش تعیین شده است، بردار ضرایب،  $\xi_i^-$  و  $\xi_i^+$  متغیرهایی هستند که قیود بالا و پایین خروجی سیستم را مشخص می‌کنند. اگر داده‌ها به صورت خطی از هم قابل جداسازی باشند،  $\xi_i^-$  و  $\xi_i^+$  سطح بهینه‌ای را تعیین می‌کنند که داده‌ها را بدون خطا و با بیشترین فاصله از صفحه و نزدیک‌ترین نقاط آموزشی (بردارهای پشتیبان) تفکیک می‌کند. مطابق شکل (۴)، اگر نقاط آموزشی به صورت  $y_i$  و بردار ورودی  $x_i$  تعریف شوند، در حالت قابل جداسازی خطی، معادله ۴ به صورت زیر قابل بیان است.

$$y = f(x) = \text{sign} \left[ \sum_{i=1}^N y_i a_i (x_i, x) + b \right] \quad (8)$$

در اینجا، y خروجی معادله است. بردار x نشان‌دهنده داده ورودی است و بردارهای  $x_i$  بردارهای پشتیبان هستند. تابع علامت<sup>۱</sup> اگر مقدار داخل عبارت آن مثبت باشد، مقدار خروجی +۱ در غیر این صورت -۱ خواهد بود.  $y_i$  برچسب<sup>۲</sup> نمونه آموزشی i است که مقدار آن +۱ یا -۱ است.  $a_i$  ضرایب لاگرانژ<sup>۳</sup> مربوط به نمونه‌های آموزشی است که در فرایند بهینه‌سازی به دست می‌آیند. این ضرایب نشان می‌دهند کدام نمونه‌ها به عنوان بردارهای پشتیبان<sup>۴</sup> عمل می‌کنند. b، بایاس یا مقدار ثابت است که در تعیین موقعیت مرز تصمیم نقش دارد. در این صورت، معادله ۸ به معادله ۹ تبدیل می‌شود. در این معادله، نمونه‌هایی که ضرایب  $a_i$  آن‌ها بزرگ‌تر از صفر هستند، به عنوان بردارهای پشتیبان شناخته می‌شوند و فقط این نمونه‌ها در تعیین مرز تصمیم مؤثرند.

<sup>4</sup> Support vectors

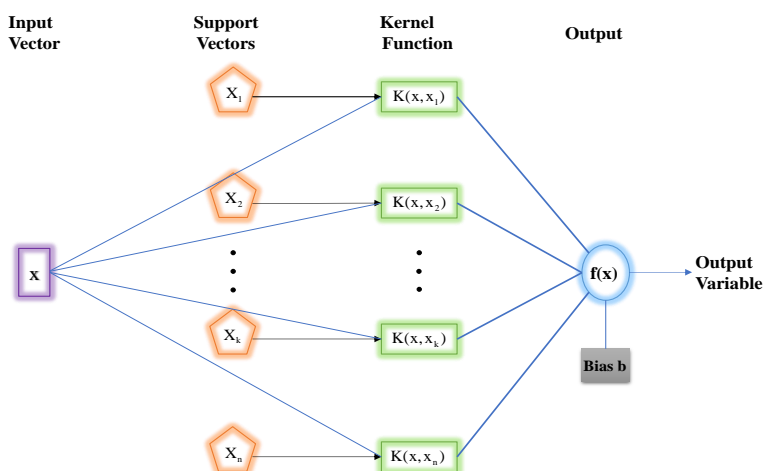
<sup>5</sup> Bagging

<sup>6</sup> Bootstrap

<sup>1</sup> Sign

<sup>2</sup> Label

<sup>3</sup> Lagrange multipliers



شکل ۴- ساختار مفهومی مدل ماشین بردار پشتیبان

Fig. 4- illustrates the concept of a Support Vector Machine (SVR)

پایان یک ناحیه داده و آغاز ناحیه دیگر را مشخص می کنند. این امر به توابع پایه (منحنی های قطعه ای رسم شده) امکان می دهد تا مدل ها را انعطاف پذیرتر کنند و به تابع خطی اجازه دهند که انحناها، آستانه ها و تغییرات دیگری را بپذیرد. ترکیب خطی توابع پایه<sup>۱</sup> (BFs) و تعاملات آن ها می تواند به شکل گیری مدل MARS کمک کند که به صورت  $f(X)$  در معادله ۱۱ نشان داده می شود (Zhang, 2020).

$$f(x) = \beta_0 + \sum_{m=1}^M \beta_m \lambda_m(x) \quad (11)$$

که در آن،  $\lambda_m$  یک تابع پایه است یا یک اسپلاین منفرد یا حاصل ضرب دو یا چند اسپلاین. در این پژوهش، برای ساده سازی هدفها فرض کردیم که حداکثر از توابع مرتبه دوم استفاده شود، هرچند اگر داده ها اجازه دهند، استفاده از مرتبه های بالاتر نیز ممکن است. روش کمترین مربعات برای برآورد  $\beta$  که ثابت است به کار می رود و توابع پایه (BFs) چندجمله ای های هموار (اسپلاین ها) با توابع قطعه ای خطی / مکعبی هستند. فرم یک تابع خطی قطعه ای مطابق معادله ۱۲ به صورت  $\max(0, x - t)$  است. این تابع به منظور سادگی به کار می رود و یک گره در مقدار  $t$  دارد.  $\max(\cdot)$  به این

در این فرآیند نمونه گیری، تقریباً یک سوم داده ها به عنوان داده های خارج از نمونه گیری در نظر گرفته می شود که برای تعیین متغیرهای مهم و برآورد بدون اریب خطا به کار می روند. پس از آن برای هر نمونه بوت استرپ، یک درخت تصمیم گسترش داده می شود و در هر مرحله تقسیم، از بین تمام  $M$  متغیر مستقل،  $m$  متغیر به طور تصادفی برای تقسیم انتخاب می شود. برای حالت رگرسیونی، نسبت  $M_m$  برابر با یک سوم است، در حالی که برای حالت طبقه بندی، مقدار  $m$  به عنوان  $m = \sqrt{M}$  پیشنهاد شده است (Siasar & Honar, 2019).

### رگرسیون اسپلاین تطبیقی چندگانه

MARS روشی است غیرپارامتریک و غیرخطی برای توسعه روابط بین مجموعه های مختلف متغیرهای ورودی و وابسته در یک فضای داده با ابعاد  $n$  که بدون نیاز به فرضیات خاصی درباره روابط عملکردی پایه ای متغیرهای ورودی و خروجی، بر اساس یک استراتژی تقسیم و غلبه عمل می کند که در آن داده های آموزشی به قطعات جداگانه با شیب های متفاوت تقسیم می شوند. گره ها، یا نقاط انتهایی قطعات،

<sup>1</sup> Basic Functions



قابل قبول به دست آید. از آنجا که انتخاب تابع پایه و موقعیت گره متغیر هر دو مبتنی بر داده و خاص مسئله هستند، MARS به عنوان یک فرآیند تطبیقی در نظر گرفته می‌شود.

معناست که مقدار آن صفر است، مگر اینکه بخش مثبت استفاده شود.

$$\max(0, x - t) = \begin{cases} x - t & x \geq t \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

### تحلیل حساسیت

برای نشان دادن تأثیر هر متغیر ورودی بر متغیر هدف، در معادله ۳ می‌توان با جابه‌جایی هر متغیر ورودی با میزان ثابت و ثابت نگه‌داشتن متغیرهای ورودی دیگر، مطابق با معادله ۱۴، از روش دامنه کسینوسی بهره برد (Momeni et al., 2014).

$$R_{ij} = \frac{\sum_{k=1}^N I_{ik} O_{jk}}{\sqrt{\sum_{k=1}^N I_{ik}^2 \cdot \sum_{j=1}^M O_{jk}^2}} \quad (14)$$

که در آن  $I_i$  و  $O_j$  به ترتیب متغیرهای ورودی و خروجی و  $N$  تعداد کل مجموعه داده‌هاست. مقدار  $R_{ij}$  بین دو مقدار صفر و یک محاسبه می‌شود که مقادیر بالاتر نشان‌دهنده رابطه قوی‌تر بین پارامترهای ورودی و هدف را نشان می‌دهد.

علاوه بر روش بالا، در تحقیق حاضر از روش آنالیز حساسیت SHAP<sup>۱</sup> استفاده گردید. مقادیر SHAP تکنیک مستقل از مدل‌سازی به منظور تفسیر پیش‌بینی‌های به دست آمده از مدل‌های یادگیری ماشین است (Lundberg, 2017). مقادیر SHAP از نظریه بازی و توضیحات محلی استفاده می‌کنند تا توزیع سهم هر ورودی را در پیش‌بینی نهایی به دست آورد (Parsa et al., 2020). به منظور حصول اطمینان از اینکه سهم هر ویژگی به‌طور عادلانه توزیع می‌شود، مقادیر SHAP از طریق معادله ۱۵ به دست می‌آید:

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n-|S|-1)!}{n!} \downarrow [v_x(S \cup i) - v_x(S)] \quad (15)$$

که در آن  $\phi_i$  نشان‌دهنده سهم ورودی  $i$  است.  $N$  مجموع همه ورودی‌های موجود در مدل است.  $n$  تعداد کل ورودی‌ها است.  $S$  زیرمجموعه‌ای از ورودی‌ها به جز ویژگی  $i$  است.

برای تولید توابع پایه (BFs)، روش MARS به صورت گام‌به‌گام میان تمام تعاملات متغیرها و تمامی موقعیت‌های ممکن گره‌های تک‌متغیره جستجو می‌کند که توسط یک الگوریتم رگرسیون تطبیقی انتخاب می‌شوند. برای ساخت مدل MARS از تکنیکی استفاده می‌شود که دارای یک فاز پیشرو برای افزودن توابع به منظور یافتن گره‌های بهبود عملکرد است، تا زمانی که بیشترین تعداد از پیش تعیین شده از عبارات شناسایی شوند و مدلی با بیش‌برازش عمدی ایجاد شود، و همچنین یک فاز پسرو که با استفاده از روش GCV (اعتبارسنجی متقاطع تعمیم‌یافته) ناکارآمدترین عبارات را حذف می‌کند تا از بیش‌برازش جلوگیری شود؛ این روش، یک آزمون برازش است که توابع پایه اضافی را جریمه می‌کند و احتمال بیش‌برازش را کاهش می‌دهد. معادله GCV برای یک مجموعه داده آموزشی با  $N$  مشاهده به صورت زیر است (Hastie et al., 2009):

$$GCV = \frac{\frac{1}{N} \sum_{i=1}^N [y_i - f(x_i)]^2}{\left[1 - \frac{M + d \times (M - 1)/2}{N}\right]^2} \quad (13)$$

که در آن  $M$  تعداد توابع پایه (BFs)،  $N$  تعداد مشاهدات،  $f(x_i)$  مقادیر تخمینی مدل MARS و  $d$  پارامتر جریمه (هزینه بهینه‌سازی هر تابع پایه برای هموارسازی فرآیند) را نشان می‌دهد. مقادیر بالاتر  $d$  به معنی تعداد کمتر گره‌ها و در نتیجه، تخمین‌های با نوسان کمتر است. فریدمن (Friedman 1991) مقدار بهینه  $2 \leq d \leq 4$  و ترجیحاً ۳ را پیشنهاد کرده است، در این پژوهش مقدار  $d = 3$  در نظر گرفته شده است. معادله ۱۳ زمانی به حداقل می‌رسد که در هر گام یک تابع پایه حذف شود تا مدلی با برازش خوب و

<sup>۱</sup> SHapley Additive exPlanations

سطح آب بالادست و نیز افزایش هد نسبی به دلیل تداخل جریان عبوری از سرریز و غیر عمود بودن حرکت راستای جریان، بر تمام سطح تاج سرریزهای کنگره‌ای، منجر به کاهش ضریب دبی شده است. به نحوی که سرریز نیم‌دایره تک سیکل در مقدار عدد نسبت هد به عمق سرریز ۰/۱۱، نسبت به سرریز با تعداد دو سیکل (N=2)، ۱۳ درصد، نسبت به سرریز با سه سیکل (N=3)، ۱۸ درصد و نسبت به سرریز با چهار سیکل (N=4)، ۲۲ درصد ضریب دبی بالاتری داشته است. می‌توان مشاهده کرد عملکرد سرریزها با افزایش سیکل‌های هارمونیک تغییر کرده است. با توجه به اثر تغییرات تعداد سیکل (N) در رابطه بین ضریب دبی و عمق نسبی، معادله‌های رگرسیون که در جدول (۲) بیان شده، برای هر سرریز با سیکل مشخص با میزان اثرگذاری شاخص ضریب تبیین محاسبه شده است.

به منظور تخمین ضریب دبی سرریزهای کنگره‌ای با پلان نیم‌دایره، در معادله ۱۸ رگرسیون غیرخطی به ازای تمامی مقادیر سیکل‌ها به شکل زیر پیشنهاد می‌شود:

$$Cd = \frac{1}{2N^{0.143}} - 0.126 \log\left(\frac{H_t}{P}\right) \quad (18)$$

معادله ۱۸ می‌تواند مقدار ضریب دبی را با شاخص  $RMSE=0.03$  و  $R^2=0.81$  هم در بخش آموزش و هم در بخش اعتبارسنجی پیش‌بینی کند.

### نتایج آنالیز حساسیت داده‌ها

در شکل (۶)، نتایج معادله ۱۴ نشان می‌دهد که متغیرهای هد آب نسبی و تعداد سیکل‌های نیم‌دایره و شعاع نسبی سرریز بیشترین تأثیر و پارامتر  $\frac{B}{P}$  کمترین تأثیر را در برآورد مقدار  $C_d$  داشته است.

$v_x(S)$  پیش‌بینی مورد انتظار با توجه به زیرمجموعه ورودی‌های  $S$  است و عبارت  $[v_x(S \cup i) - v_x(S)]$  نشان‌دهنده اختلاف ناشی از اضافه کردن ورودی  $i$  است.

### معیارهای ارزیابی

به منظور ارزیابی مدل‌های مورداستفاده در این تحقیق از شاخص‌های آماری ضریب تعیین، ریشه میانگین مربعات خطا به ترتیب در معادله‌های ۱۶ و ۱۷ استفاده گردید:

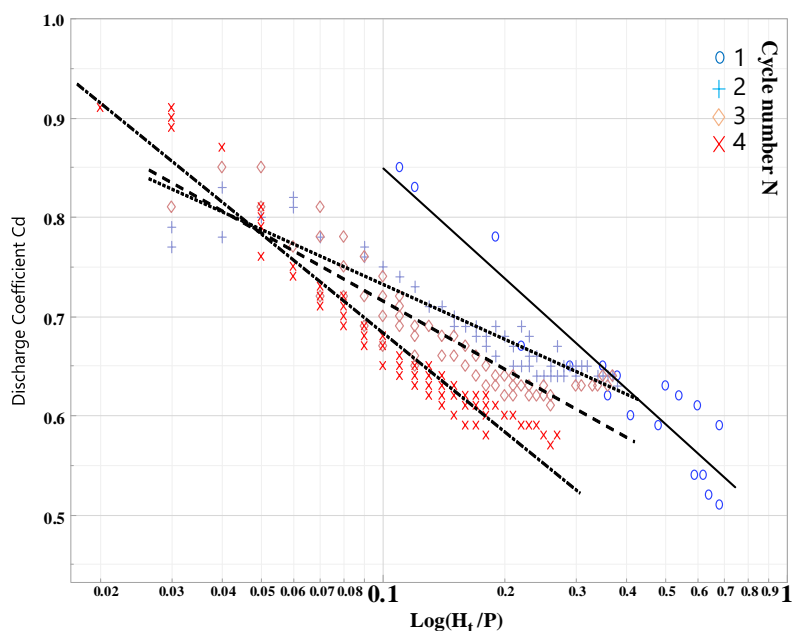
$$R^2 = \frac{[\sum_{i=1}^N (O_i - \bar{O}_i)(P_i - \bar{P}_i)]^2}{\sum_{i=1}^N (O_i - \bar{O}_i)^2 \sum_{i=1}^N (P_i - \bar{P}_i)^2} \quad (16)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (O_i - P_i)^2}{N}} \quad (17)$$

در معادله‌های بالا،  $O_i$  مقادیر ضریب دبی مشاهداتی (واقعی)،  $P_i$  مقادیر ضریب دبی پیش‌بینی شده،  $\bar{O}_i$  میانگین مقادیر ضریب دبی مشاهداتی،  $\bar{P}$  میانگین مقادیر ضریب دبی پیش‌بینی شده و  $N$  تعداد مشاهدات است. لازم استگفته شود مقادیر  $R^2$  نزدیک یک و مقادیر  $RMSE$  نزدیک به صفر، نشان‌دهنده دقت بیشتر و عملکرد بهتر مدل هستند.

### نتایج و بحث

به منظور ارزیابی نتایج مدل‌های آزمایشگاهی جمع‌آوری شده، پارامتر بدون بعد ضریب دبی نسبت به عمق نسبی به ازای تعداد سیکل‌های سرریز کنگره‌ای در شکل (۵) نشان داده شده است. مشاهده می‌شود که عملکرد سرریزهای کنگره‌ای هارمونیک با افزایش تعداد سیکل بهبود پیدا کرده است به گونه‌ای که در مقادیر هد نسبی پایین‌تر از ۰/۰۵، میزان ضریب دبی عبوری سرریز با ۴ سیکل نیم‌دایره (N=4)، بیشتر از سائز سرریزها بوده است. این روند به ازای افزایش



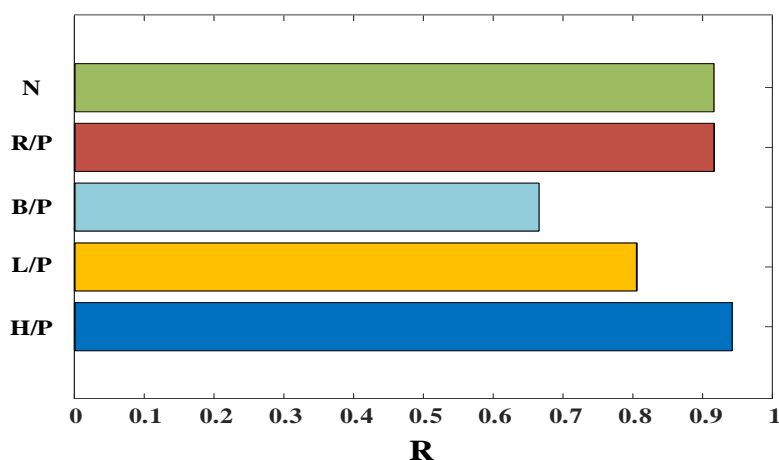
شکل ۵- تغییرات ضریب دبی به نسبت عمق به ارتفاع سرریز

Fig. 5- Relation to the depth to the height of the weirs with discharge coefficient

جدول ۲- معادله‌های رگرسیون غیرخطی برای هر سرریز با سیکل مشخص

Table 2- Nonlinear regression equations for each Weirs with specified cycle

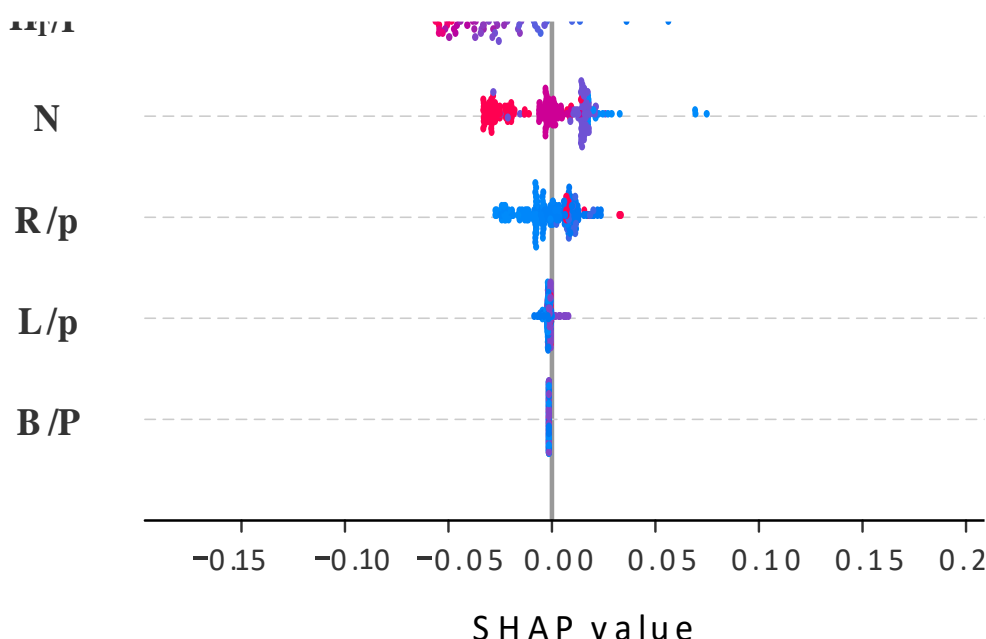
$R^2$	رگرسیون	تعداد سیکل
۰/۸۹	$Cd = 0.4802 - 0.369 \log\left(\frac{H_t}{P}\right)$	N=1
۰/۹	$Cd = 0.5482 - 0.1838 \log\left(\frac{H_t}{P}\right)$	N=2
۰/۸۳	$Cd = 0.4872 - 0.228 \log\left(\frac{H_t}{P}\right)$	N=3
۰/۹۴	$Cd = 0.352 - 0.331 \log\left(\frac{H_t}{P}\right)$	N=4



شکل ۶- میزان تأثیر مولفه‌های مختلف ورودی به مدل در پیش‌بینی ضریب دبی

Fig. 6 - The impact of different input components to model in predicting Cd

نمودار شکل (۷) مقدار مولفه SHAP را برای هر ورودی به مدل نشان می‌دهد. هر نقطه رنگی روی نمودار نمایانگر مقدار خاص از یک ورودی در مجموعه داده است، در حالی که محور  $y$  و محور  $x$  به ترتیب نمایانگر ورودی‌ها به مدل و مقادیر SHAP هستند که میزان تأثیر آن را نشان می‌دهند. طیف رنگی اختصاص داده شده به هر نقطه داده، مقادیر متغیر ویژگی‌ها را منعکس می‌کند. به طور دقیق‌تر، آبی نشان‌دهنده مقادیر کمتر و قرمز نشان‌دهنده مقادیر بالاتر است. مشابه تحلیل روش دامنه کسینوسی، مقادیر SHAP نشان می‌دهند که متغیر هد آب نسبی بیشترین تأثیر را بر پیش‌بینی ضریب دبی در سرریزهای هارمونیک دارد. علاوه بر این، مقدار SHAP برای تعداد سیکل‌های سرریز به عنوان دومین ویژگی مهم شناسایی شد.



شکل ۷- نتایج خروجی مدل آنالیز حساسیت SHAP  
 Fig. 7- Results of a feature importance in SHAP analysis

بر اساس شکل (۷)، مقادیر هد نسبی بیشتر مقدار SHAP منفی داشتند که نشان‌دهنده تأثیر منفی بر پیش‌بینی‌های نهایی ضریب دبی است. از سوی دیگر، مقادیر SHAP برای طول و عرض نسبی سرریز تقریباً صفر بودند که نشان می‌دهد این ویژگی‌ها تأثیر قابل توجهی بر پیش‌بینی‌های ضریب دبی نداشتند. علاوه بر این، مقادیر SHAP برای شعاع نسبی سرریز اندکی مثبت بود که نشان‌دهنده تأثیر مثبت است.

برای بهبود دقت پیش‌بینی، پنج ترکیب مختلف ورودی بر اساس اهمیت هر پارامتر ورودی به دست آمده از آنالیز بخش پیشین طراحی شد تا با اجرای مدل‌های یادگیری ماشین، مؤثرترین ترکیب شناسایی شود. با توجه به تأثیر بالای عمق نسبی بر میزان ضریب دبی، به‌منظور تعریف سناریوهای ترکیب ورودی‌های مختلف این مؤلفه در تمامی معادله‌ها دیده شده است. در جدول (۳)، ترکیب‌های خاص ورودی برای پیش‌بینی ضریب دبی نشان داده شده است.

جدول ۳- سناریوهای مختلف ورودی به منظور مدل‌سازی ضریب دبی

Table 3- Different input scenarios to model the discharge coefficient

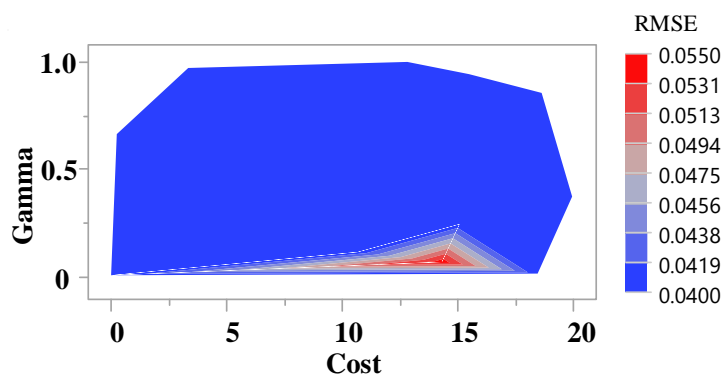
مدل	رگرسیون
Model 1	$Cd = f\left(\frac{H_t}{P}\right)$
Model 2	$Cd = f\left(\frac{H_t}{P}, N\right)$
Model3	$Cd = f\left(\frac{H_T}{P}, \frac{R}{P}, N\right)$
Model4	$Cd = f\left(\frac{H_T}{P}, \frac{R}{P}, \frac{L}{P}, N\right)$
Model5	$Cd = f\left(\frac{H_T}{P}, \frac{R}{P}, \frac{B}{P}, \frac{L}{P}, N\right)$

عملکرد مدل SVR برای تخمین دبی جریان بررسی شد. (Cost) ۰/۹۵۸، مشابه آنچه در شکل (۸) نیز آورده شده به‌منظور توسعه مدل SVR توابع کرنل مختلفی که در است نسبت به سایر مدل‌های دیگر دقت بهتری دارد و می‌تواند به عنوان مدل بهینه شده ماشین بردار پشتیبان استفاده بیان شده داده در جدول (۴) نشان می‌دهد که تابع شعاعی با پارامترهای ضریب پنالتی (Gamma) ۰/۰۱۴ و پارامتر خطا

جدول ۴- بررسی توابع مختلف به عنوان کرنل در مدل SVR

Table 4- Examination of different functions as kernel in SVR model

مدل	تابع کرنل	Cost	Gamma	Training RMSE	Test	
					RMSE	R <sup>2</sup>
1	Linear	۰/۶۵۰	۰/۰۴۰	۰/۰۴۳	.	۹/۹۵۶
2	Linear	۰/۶۵۰	۰/۰۴۰	۰/۰۴۳	.	۹/۰۰۹
3	Linear	۰/۶۵۰	۰/۰۴۰	۰/۰۴۳	.	۹/۶۴۲
4	Linear	۰/۶۵۱	۰/۰۴۰	۰/۰۴۳	.	۰/۳۱۵
5	Radial Basis Function	۰/۹۵۵	۰/۰۱۴	۰/۰۱۱	۰/۹۴۰	۱۵/۴۹۷
6	Radial Basis Function	۰/۹۵۶	۰/۰۱۴	۰/۰۱۱	۰/۹۹۷	۱۲/۸۲۶
7	Radial Basis Function	۰/۹۵۶	۰/۰۱۴	۰/۰۱۳	۰/۶۹۹	۹/۲۲۷
8	Radial Basis Function	۰/۹۵۷	۰/۰۱۴	۰/۰۱۲	۰/۸۰۶	۱۱/۷۱۶
9	Radial Basis Function	۰/۹۵۸	۰/۰۱۴	۰/۰۱۲	۰/۶۷۳	۱۷/۶۶۹

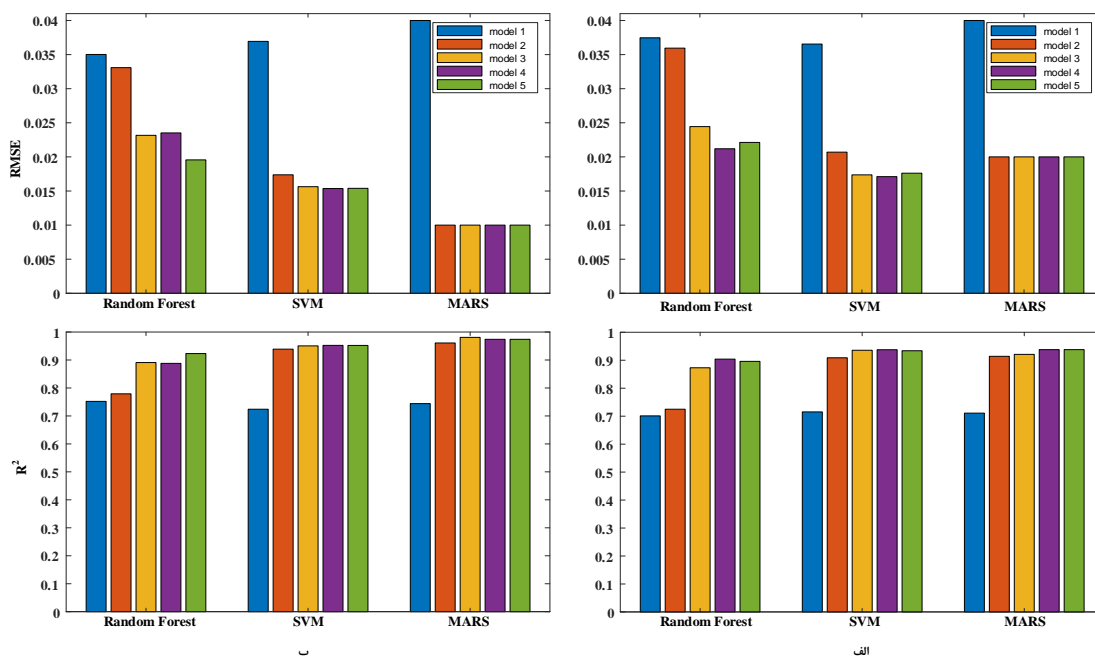


شکل ۸- تغییرات پارامترهای Cost و Gamma مدل SVR در دقت تخمین میزان ضریب دبی

Fig. 8- Cost and Gamma variations in the SVR Model to Estimation of Discharge Coefficient

پیش‌بینی‌شده توسط ترکیب ورودی چهار تحلیل شده است. همچنین عملکرد پیش‌بینی مدل‌های یادگیری ماشین برای سرریزهای هارمونیک مختلف به صورت نمودار پراکندگی در شکل (۱۰) برای تمامی سناریوهای ترکیب ورودی آورده شده است. مشاهده می‌شود که مدل‌های SVR و MARS نسبت به مدل جنگل تصادفی RF عملکرد پیش‌بینی خوبی را در هر دو مجموعه داده آموزشی و آزمایشی نشان می‌دهند.

عملکرد ترکیب‌های ورودی جدول (۳) با استفاده از دو شاخص RMSE و  $R^2$  که در شکل (۹) نشان داده شده، ارزیابی شده است. نتایج بر مبنای داده‌های بخش اعتبار سنجی (شکل ۹ الف) نشان می‌دهد که ترکیب ورودی مدل شماره چهار در بین مدل‌های SVR و MARS بهترین عملکرد کلی را دارد و پس از آن ترکیب‌های ورودی ۵، ۳، ۲ و ۱ قرار می‌گیرند. بنابراین، عملکرد مدل‌های مختلف یادگیری ماشین در بخش‌های بعدی بر اساس نتایج

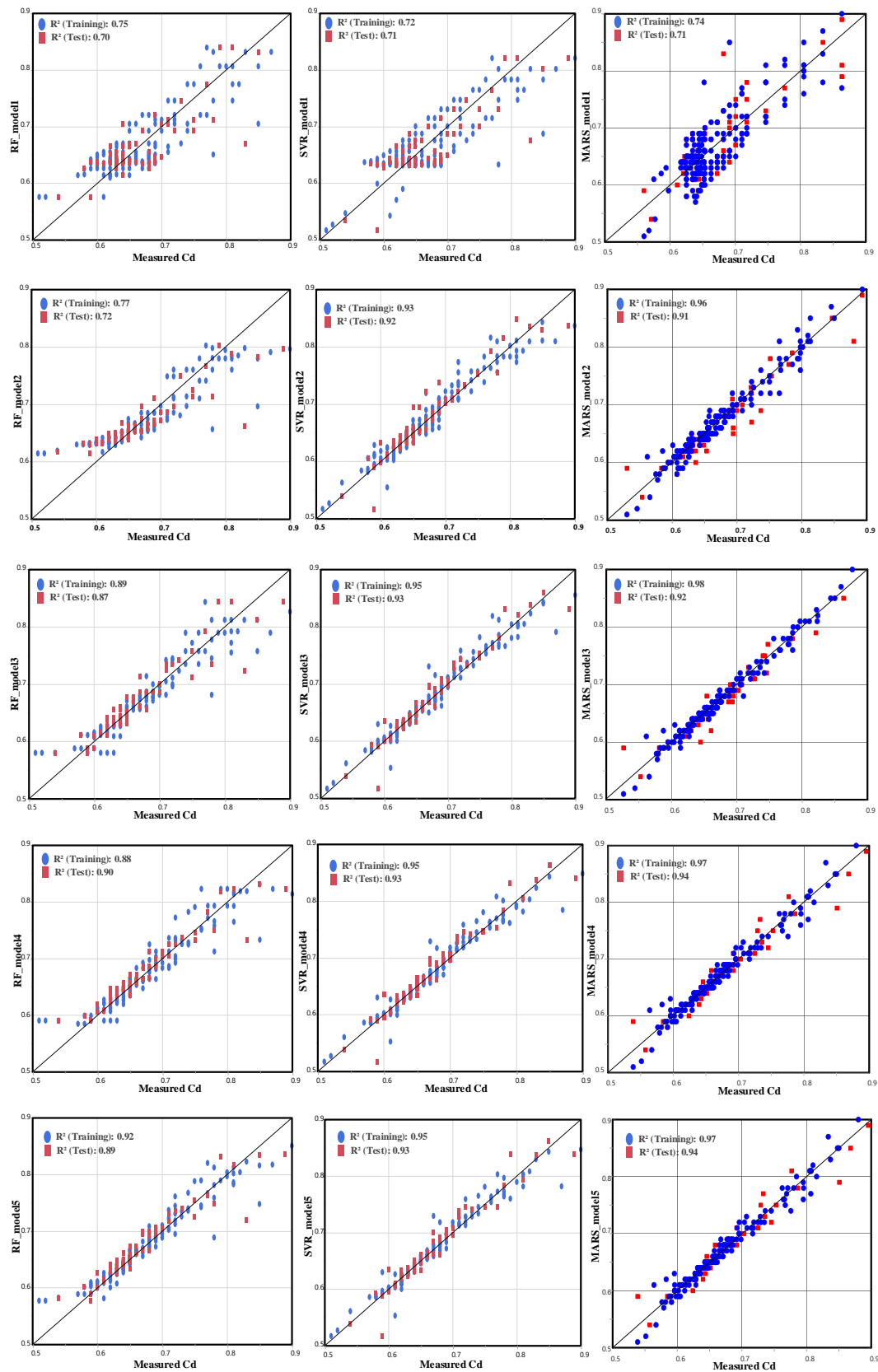


شکل ۹- تأثیر ورودی‌های مختلف در پیش‌بینی ضریب دبی با استفاده از مدل‌های توسعه یافته در بخش الف) اعتبار سنجی و ب) در بخش آموزش

Fig. 9- The impact of different inputs in predicting the discharge coefficient using the models developed in the A) Testing section and B) in the Training section



کاربرد هوش مصنوعی در پیش‌بینی ضریب دبی در سرریزهای کنگره‌ای با پلان هارمونیک



شکل ۱۰- نمودار پراکندگی ضرایب دبی اندازه‌گیری شده نسبت به مدل‌های MARS، SVR و RF

Fig. 10- Scatter plot of the measured discharge coefficients compared to MARS, SVR and RF models

در فرآیند توسعه مدل MARS، در مرحله اول ۲۱ تابع پایه در نظر گرفته شد و در مرحله دوم (مرحله هرس کردن)، یک تابع پایه حذف گردید. در نهایت، یک مد MARS بهینه با ۲۰ تابع پایه حاصل شد. فرم کلی مدل MARS به دست آمده در معادله ۱۹ ارائه شده است و فرم توسعه یافته آن در جدول (۵) آمده است. استفاده از معادلات حاصل از توسعه مدل MARS برای هدف های عملی آسان است و می توان ضریب دبی به دست آمده از این معادله را از طریق محاسبات دستی انجام داد.

$$C_d = 0.6928 + \sum_{M=1}^{20} \beta_M h_M(x) \quad (19)$$

برای پیش بینی ضریب دبی، بهترین عملکرد مدل درخت تصمیم در الگوریتم جنگل تصادفی با شاخص های  $R^2 = 0.902$ ,  $RMSE = 0.02$  بر مبنای ترکیب ورودی شماره چهارم در شکل (۱۱) نشان داده شده است.

جدول ۵- ضرایب تابع پایه نشان داده شده در معادله ۱۹ به منظور پیش بینی ضریب دبی

Table 5- The coefficients of the base function shown in equation (19) to predict the discharge coefficient

$(\beta_M)$	معادله	تابع پایه
-0.7531	BF1 = max(0, $H_T/P - 0.13$ )	$h_1(x)$
0.8274	BF2 = max(0, $0.13 - H_T/P$ )	$h_1(x)$
0.9347	BF3 = max(0, $N - 3$ ) * max(0, $H_T/P - 0.08$ )	$h_1(x)$
1.7096	BF4 = max(0, $N - 3$ ) * max(0, $0.08 - H_T/P$ )	$h_1(x)$
-0.0172	BF5 = max(0, $R/P - 0.5$ )	$h_1(x)$
1.2601	BF6 = BF2 * max(0, $N - 2$ )	$h_1(x)$
-10.7415	BF7 = BF2 * max(0, $2 - N$ )	$h_1(x)$
-2.0767	BF8 = BF1 * max(0, $N - 3$ )	$h_1(x)$
0.5215	BF9 = BF1 * max(0, $3 - N$ )	$h_1(x)$
2.1012	BF10 = max(0, $0.22 - H_T/P$ )	$h_1(x)$
-3.1755	BF11 = BF10 * max(0, $0.75 - R/P$ )	$h_1(x)$
0.0206	BF12 = max(0, $N - 3$ ) * max(0, $B/P - 2$ )	$h_1(x)$
-22114.5081	BF13 = BF10 * max(0, $R/P - 0.5$ )	$h_1(x)$
-6.6756	BF14 = BF10 * max(0, $0.5 - R/P$ )	$h_1(x)$
22117.8837	BF15 = BF10 * max(0, $R/P - 0.5$ )	$h_1(x)$
-1.4013	BF16 = BF10 * max(0, $B/P - 2$ )	$h_1(x)$
0.9593	BF17 = BF10 * max(0, $2 - B/P$ )	$h_1(x)$
-0.2331	BF18 = max(0, $3 - N$ ) * max(0, $0.45 - R/P$ )	$h_1(x)$
0.8107	BF19 = max(0, $H_T/P - 0.22$ ) * max(0, $N - 2$ )	$h_1(x)$
-0.5991	BF20 = max(0, $H_T/P - 0.22$ ) * max(0, $2 - N$ )	$h_1(x)$

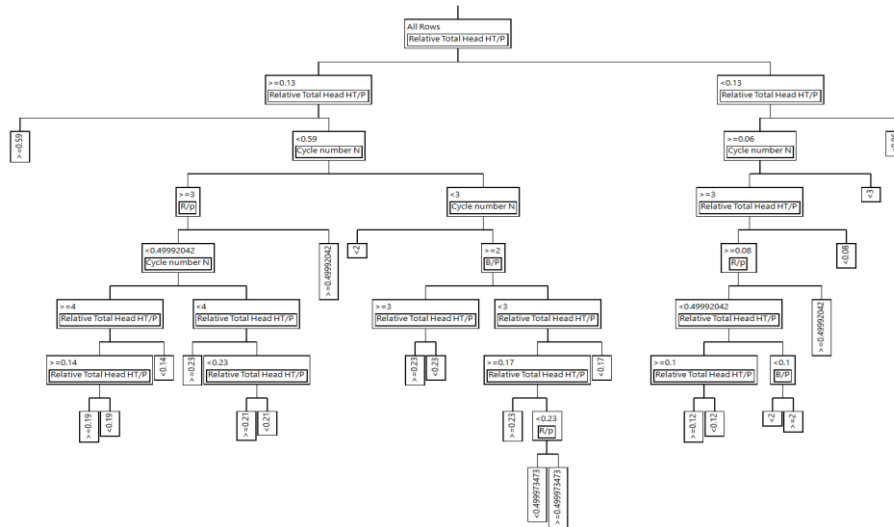
$$C_d = 0.581 \left(\frac{H_T}{P}\right)^{-0.121} \left(\frac{B}{P}\right)^{-0.035} N^{-0.047} \quad (20)$$

ارهام نمازی و مظفری (Arham Namazi & Mozaffari, 2023) برای حالت سرریز با تاج تک نیم دایره معادله زیر را برای برآورد ضریب دبی تعریف کرده اند:

$$C_d = 0.4731 - 0.158 \ln\left(\frac{H_t}{P}\right) \quad (21)$$

پیش تر گفته شد که در زمینه مدل های هوش مصنوعی در پیش بینی ضریب دبی سرریزهای کنگره ای تا کنون مطالعه ای گزارش نشده است. یلدیز و همکاران (Yıldız et al., 2024) معادله ای رگرسیون غیرخطی زیر را به منظور تخمین ضریب دبی سرریزهای کنگره ای با پلان هارمونیک به شکل زیر پیشنهاد داده اند:

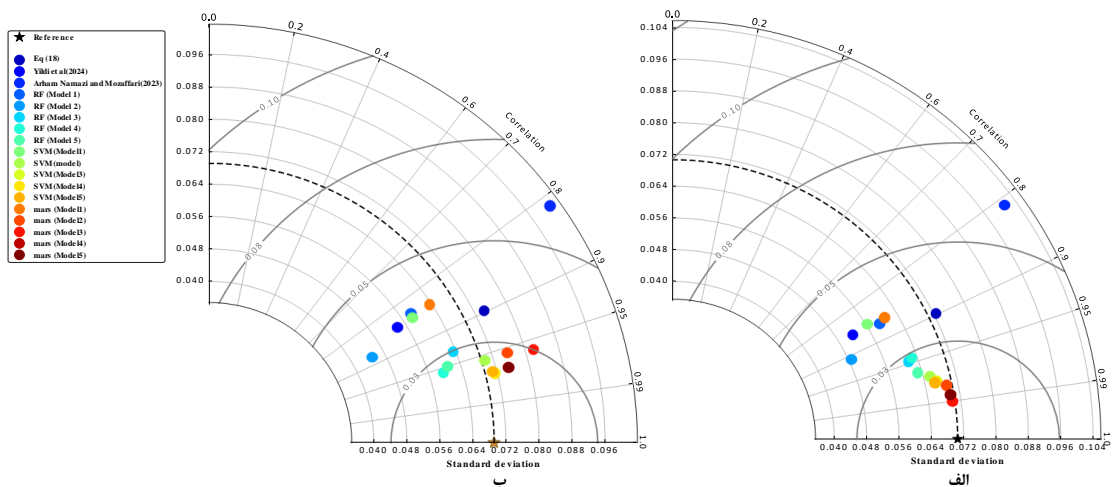
کاربرد هوش مصنوعی در پیش‌بینی ضریب دبی در سرریزهای کنگره‌ای با پلان هارمونیک



شکل ۱۱- ساختار بهترین درخت تصمیم در مدل RF به منظور برآورد ضریب دبی  
 Fig. 11 - The structure of the Best Decision tree in the RF model to estimate the Cd

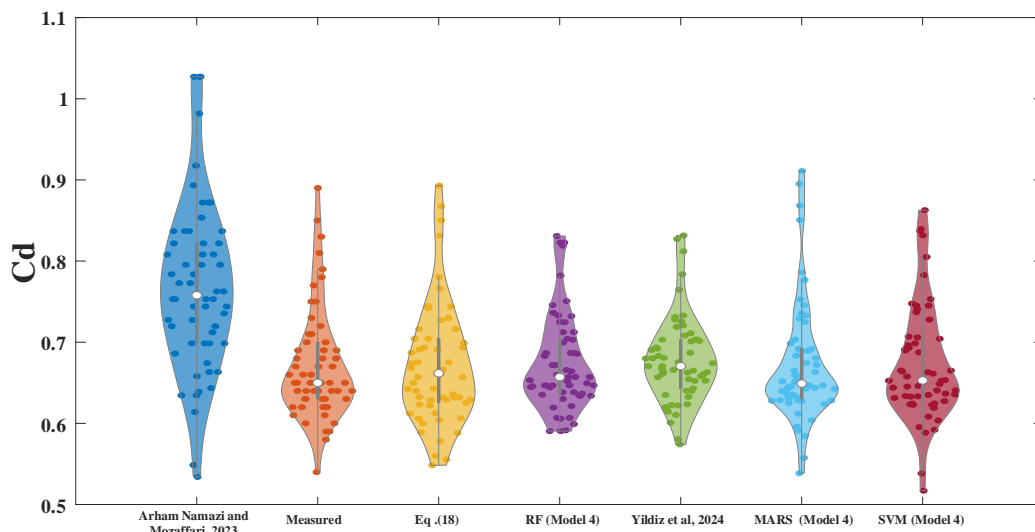
ضریب دبی نسبت به روش‌های کلاسیک و ناپارامتری بیشتر تأیید می‌شود. مشاهده می‌شود که نقاط محاسبه‌شده توسط الگوریتم جنگل تصادفی (RF) برای هر نوع سیکل سرریز هارمونیک، در مقایسه با نقاط ارائه‌شده توسط الگوریتم‌های دیگر، فاصله بیشتری از نقطه هدف دارند. علاوه بر این، معادله غیرخطی ارائه شده در این تحقیق نیز (معادله ۱۸) نسبت به مدل‌های رگرسیون سنتی گزارش داده شده، عملکرد بهتری داشته است.

شکل (۱۲) نمودارهای تیلور را برای داده‌های اندازه‌گیری شده و داده‌های پیش‌بینی شده (شامل مجموعه‌های آموزش و آزمایش) نمایش می‌دهد که توسط مدل‌های مختلف یادگیری ماشین پیش‌بینی شده‌اند. همان‌گونه که در شکل ۱۲-ب مشاهده می‌شود، فاصله نقاطی که توسط الگوریتم‌های SVR و MARS به دست آمده‌اند، هر دو نزدیک‌ترین به نقطه اندازه‌گیری شده مرجع هستند. از این رو می‌توان نتیجه گرفت عملکرد الگوریتم بهینه‌سازی شده ماشین بردار پشتیبان و الگوریتم MARS در تخمین میزان



شکل ۱۲- دیاگرام تیلور عملکرد مدل‌های توسعه‌یافته (الف) بخش آموزش (ب) بخش اعتبارسنجی  
 Fig. 12 - Taylor diagram of the performance of the developed models

شکل (۱۳) نمودارهای ویولنی را برای داده‌های اندازه‌گیری شده و داده‌های پیش‌بینی شده نشان می‌دهد که توسط مدل‌های مختلف یادگیری ماشین استنتاج شده‌اند. در این شکل می‌توان مشاهده کرد که هر سه مدل SVR، RF و MARS اساساً چارک اول، سوم و میانه مشابهی را با داده‌های اندازه‌گیری شده پیش‌بینی می‌کنند، در حالی که چارک اول یا سوم در معادله ۲۱ و معادله ۱۸ تفاوت قابل توجهی با مقادیر اندازه‌گیری شده دارد. علاوه بر این، از دیدگاه توزیع کلی داده‌ها، الگوریتم SVR و MARS، نسبت به الگوریتم RF، به ترتیب شکل مشابه‌تری با داده‌های اندازه‌گیری شده نشان می‌دهند، که این امر قابلیت پیش‌بینی قابل قبول استراتژی ماشین بردار پشتیبان را نشان می‌دهد.



شکل ۱۳- نمودار ویولنی مقایسه مدل‌های یادگیری ماشین با معادله‌های پیشین

Figure 13- Violin diagram comparing machine learning models with previous relationships

MARS را تأیید می‌کند. نتایج پیش‌بینی این مطالعه به شرح

زیر خلاصه شده است:

برای کاهش خطاهای پیش‌بینی مدل‌های یادگیری ماشین، حساسیت پارامترها ابتدا تحلیل شد تا اهمیت نسبی پارامترهای مختلف ورودی مشخص شود. بر این اساس، پنج ترکیب ورودی مختلف طراحی و به مدل‌های یادگیری ماشین اعمال شد. مقایسه شاخص‌های آماری بین نتایج پیش‌بینی شده و داده‌های آزمایشگاهی نشان داد که برای پیش‌بینی ضریب دبی این نوع از سرریزها، ترکیب ورودی مشترک بهینه شامل  $N, \frac{H_T}{P}, \frac{R}{P}, \frac{L}{P}$  است. با استفاده از بهترین ترکیب ورودی، نتایج نشان داد که الگوریتم‌های SVR و MARS از مدل‌های مبتنی بر درخت تصمیم، مانند جنگل

### نتیجه‌گیری

سرریزهای کنگره‌ای همواره به عنوان راه‌حلی مؤثر برای افزایش راندمان دبی سیلاب در مواقعی پیشنهاد می‌شوند که فضای محدودی برای ساخت سرریز وجود دارد. سرریزهای کنگره‌ای طول بیشتری نسبت به سرریزهای ساده خطی دارند. بنابراین سیلاب با عمق کمتری از روی آن عبور خواهد کرد. به دلیل ارتباط پیچیده بین ضریب دبی و پارامترهای مرتبط، معادلات تجربی به‌سختی می‌توانند ضریب دبی را با دقت قابل قبول پیش‌بینی کنند. در مطالعه حاضر، سه مدل مختلف یادگیری ماشین برای مدل‌سازی ضریب دبی سرریزهای نیم‌دایره با طرح هارمونیک توسعه داده شده است که مزایای الگوریتم ماشین بردار پشتیبان (SVR) مدل

تصادفی (RF) در پیش‌بینی برای هر نوع از سرریز هارمونیک مطالعات پیشین استفاده شده‌اند نشان داد که این مدل نتایج با سیکل‌های مختلف عملکرد بهتری دارند. مقایسه مدل رضایت‌بخشی ارائه می‌دهد و در دقت، عملکرد بهتری نسبت MARS با دیگر مدل‌های رگرسیونی غیر خطی که در به آن‌ها دارد.

## سپاسگزاری

حمایت‌های مالی معاونت پژوهش و فناوری دانشگاه شهید چمران اهواز در قالب پژوهانه (SCU.WH1403.43525) زمینه اجرای این تحقیق را فراهم آورد که بدین‌وسیله از ایشان قدردانی می‌شود.

## مراجع

- Arham Namazi, F. S., & Mozaffari, J. (2023). Investigation of labyrinth weirs discharge coefficient with the same length. *Flow Measurement and Instrumentation*, 94, 102468. <https://doi.org/10.1016/j.flowmeasinst.2023.102468>
- Behdarvandi, B., Zayeri, M. R., Ghomeshi, M., & Daryaei, M. (2024). Performance of Support Vector Machines, Random Forest, and GMDH Methods for Predicting Flow Rate from Rectangular Flap Gates. *Irrigation and Water Engineering*. <https://doi.org/10.22125/iwe.2024.460044.1811>
- Bilhan, O., Aydin, M. C., Emiroglu, M. E., & Miller, C. J. (2018). Experimental and CFD Analysis of Circular Labyrinth Weirs. *Journal of Irrigation and Drainage Engineering*, 144(6), 04018007. [https://doi.org/10.1061/\(ASCE\)IR.1943-4774.0001301](https://doi.org/10.1061/(ASCE)IR.1943-4774.0001301)
- Bonakdari, H., Ebtahaj, I., Gharabaghi, B., Sharifi, A., & Mosavi, A. (2020). *Prediction of Discharge Capacity of Labyrinth Weir with Gene Expression Programming* (No. arXiv:2002.02751). arXiv. <https://doi.org/10.48550/arXiv.2002.02751>
- Borghei, S. M., & Parvaneh, A. (2011). Discharge characteristics of a modified oblique side weir in subcritical flow. *Flow Measurement and Instrumentation*, 22(5), 370–376. <https://doi.org/10.1016/j.flowmeasinst.2011.04.009>
- Borisova, N. M., & Belikov, V. V. (2023). Optimization of Spillway Design of Reconstructed Shavan Dam Using Numerical 2D-Modeling. *Power Technology and Engineering*, 57(1), 67–74. Scopus. <https://doi.org/10.1007/s10749-023-01624-7>
- Dutta, D., Mandal, A., & Afzal, M. S. (2020). Discharge performance of plan view of multi-cycle W-form and circular arc labyrinth weir using machine learning. *Flow Measurement and Instrumentation*, 73, 101740. <https://doi.org/10.1016/j.flowmeasinst.2020.101740>
- Emami, S., Parsa, J., Emami, H., & Abbaspour, A. (2021). An ISaDE algorithm combined with support vector regression for estimating discharge coefficient of W-planform weirs. *Water Supply*, 21(7), 3459–3476.
- Friedman, J. H. (1991). Multivariate adaptive regression splines. *The Annals of Statistics*, 19(1), 1–67.
- Ghaderi, D., Ebrahimnezhadian, H., & Mollazadeh, M. (2024). Three-dimensional analysis of the performance of circular stepped spillways in the skimming flow regime. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 46(7). Scopus. <https://doi.org/10.1007/s40430-024-05004-8>
- Hagbin, M., Sharafati, A., Aghamajidi, R., Asadollah, S. B. H. S., Noghani, M. H. M., & Jalón, M. L. (2022). Determination of discharge coefficient of stepped morning glory spillway using a hybrid data-driven method. *Flow Measurement and Instrumentation*, 85, 102161.
- Hastie, T., Tibshirani, R., Friedman, J. H., & Friedman, J. H. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (Vol. 2). Springer. <https://link.springer.com/book/10.1007/978-0-387-21606-5>

- Hussein, B. S., & Jalil, S. A. (2024). Influence of Labyrinth Side Weir Shape Modification on the Hydrodynamic Performance: Experimental and Numerical Study. *Arabian Journal for Science and Engineering*. Scopus. <https://doi.org/10.1007/s13369-024-09563-9>
- Ikinciogullari, E. (2024). Energy dissipation performance of labyrinth and harmonic stepped spillways. *Journal of Hydroinformatics*, 26(10), 2668–2682. Scopus. <https://doi.org/10.2166/hydro.2024.221>
- Kisi, O., Emiroglu, M. E., Bilhan, O., & Guven, A. (2012). Prediction of lateral outflow over triangular labyrinth side weirs under subcritical conditions using soft computing approaches. *Expert Systems with Applications*, 39(3), 3454–3460.
- Li, S., Shen, G., Parsaie, A., Li, G., & Cao, D. (2024). Discharge modeling and characteristic analysis of semi-circular side weir based on the soft computing method. *Journal of Hydroinformatics*, 26(1), 175–188.
- Lundberg, S. (2017). A unified approach to interpreting model predictions. *arXiv Preprint arXiv:1705.07874*.
- Momeni, E., Nazir, R., Jahed Armaghani, D., & Maizir, H. (2014). Prediction of pile bearing capacity using a hybrid genetic algorithm-based ANN. *Measurement*, 57, 122–131. <https://doi.org/10.1016/j.measurement.2014.08.007>
- Parsa, A. B., Movahedi, A., Taghipour, H., Derrible, S., & Mohammadian, A. K. (2020). Toward safer highways, application of XGBoost and SHAP for real-time accident detection and feature analysis. *Accident Analysis & Prevention*, 136, 105405. <https://www.sciencedirect.com/science/article/pii/S0001457519311790>
- Rao, S. S., & Shukla, M. K. (1971). Characteristics of flow over weirs of finite crest width. *Journal of the Hydraulics Division*, 97(11), 1807–1816. <https://doi.org/10.1061/JYCEAJ.0003138>
- Safarrazavi Zadeh, M., Esmaeili Varaki, M., & Biabani, R. (2021). Experimental study on flow over sinusoidal and semicircular labyrinth weirs. *ISH Journal of Hydraulic Engineering*, 27(sup1), 304–313. <https://doi.org/10.1080/09715010.2019.1644679>
- Sangsefidi, Y., Mehraein, M., & Ghodsian, M. (2018). Experimental study on flow over in-reservoir arced labyrinth weirs. *Flow Measurement and Instrumentation*, 59, 215–224. <https://www.sciencedirect.com/science/article/pii/S0955598616303156>
- Shafiei, S., Najarchi, M., & Shabanlou, S. (2020a). A novel approach using CFD and neuro-fuzzy-firefly algorithm in predicting labyrinth weir discharge coefficient. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 42(1), 44. <https://doi.org/10.1007/s40430-019-2109-9>
- Shafiei, S., Najarchi, M., & Shabanlou, S. (2020b). Simulation of labyrinth weir discharge coefficient by modern artificial intelligence models. *Modares Civil Engineering Journal*, 20(1), 204–218.
- Siasar, H., & Honar, T. (2019). Application of support vector machine, CHAID and random forest models, in estimated daily reference evapotranspiration in northern Sistan and Baluchestan Province. *Iranian Journal of Irrigation & Drainage*, 13(2), 378–388.
- Simsek, O., Gumus, V., & Ozluk, A. (2023). Prediction of discharge coefficient of the trapezoidal broad-crested weir flow using soft computing techniques. *Neural Computing and Applications*, 35(24), 17485–17499.
- Suleimany, J. M. F. S., Aurahman, T. H., & S.mamand, B. (2022). Flow simulation over semicircular labyrinth weir using ANSYS-fluent. *Tikrit Journal of Engineering Sciences*, 29(1), 83–98. Scopus. <https://doi.org/10.25130/tjes.29.1.7>
- Yıldız, A., Marti, A. İ., & Göğüş, M. (2024). The hydraulic investigation of harmonic plan weirs. *Flow Measurement and Instrumentation*, 95, 102512. <https://doi.org/10.1016/j.flowmeasinst.2023.102512>
- Zare, H., Vaghefi, M., Mahmoudi, A., & Behroozi, A. M. (2023). Experimental Exploration of Flow Hydraulics and Discharge Coefficient for an Inclined Circular Labyrinth Weir. *Water Resources Management*, 37(11), 4521–4536. Scopus. <https://doi.org/10.1007/s11269-023-03577-y>
- Zayeri, M. (2023). Discharge Prediction in Flumes with Trapezoidal Contraction by Machine Learning Techniques. *Irrigation and Drainage Structures Engineering Research*, 24(90), 55–70.
- Zhang, W. (2020). *MARS Applications in Geotechnical Engineering Systems: Multi-Dimension with Big Data*. Springer Singapore. <https://doi.org/10.1007/978-981-13-7422-7>



**Original Research**

## **Application of Artificial Intelligence in Predicting the Discharge Coefficient of Labyrinth Weirs with Harmonic Plans**

**A. Savaedi, M. Zayeri\*, M. Ghomeshi, M. Daryaei**

**\*Corresponding Author:** Assistant Professor, Department of Hydraulic Structures, Faculty of Water and Environmental Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Iran

**Received:** 19 December 2024, **Accepted:** 4 January 2025

**Email:** m.zayri@scu.ac.ir

**https://doi.org/** 10.22092/IDSER.2025.368037.1602

### **Extended Abstract**

#### **Introduction**

Dams, as barriers constructed across rivers, are comprised of essential components such as the body, spillway, and drainage systems. Various labyrinth spillway designs, including triangular, trapezoidal, circular, and polygonal horizontal layouts, extend the effective flow path over a fixed width compared to linear spillways. Researchers aim to identify optimal designs balancing high performance and cost-efficiency. Recent advancements highlight the integration of optimization methods and computational fluid dynamics (CFD) to improve labyrinth spillway designs. Studies have explored the hydraulic and geometric factors affecting discharge coefficients ( $C_d$ ) and flow velocity. Research includes the application of artificial intelligence (AI) models such as artificial neural networks (ANNs), adaptive neuro-fuzzy inference systems (ANFIS), and regression techniques to predict  $C_d$ . Notable contributions demonstrate that AI models effectively capture complex nonlinear relationships between geometric parameters and flow rates, outperforming traditional methods. For instance, models like support vector machines (SVM) and adaptive regression spline (MARS) have demonstrated high accuracy in predicting  $C_d$ .

Despite advancements, precise predictive models for labyrinth spillways with harmonic plans remain underdeveloped. This study addresses this gap by introducing new methodologies, including SVM, random forests (RF), and MARS, to predict  $C_d$ . It also quantifies the influence of dimensionless parameters on  $C_d$ , synthesizing experimental data to enhance understanding and bridge existing research gaps.

#### **Methodology**

In this study, soft computing models were developed using experimental results from Arham Namazi and Mozaffari (2023) and Yıldız *et al.* (2024). To evaluate the accuracy of proposed soft computing equations in estimating the discharge coefficient ( $C_d$ ) for circular labyrinth weirs arranged harmonically in open channels, the following experimental data were utilized: Yıldız *et al.* (2024): conducted 215 experiments for weirs with three different heights ( $P = 20$  cm,  $P = 30$  cm, and  $P = 40$  cm) and three different cycle numbers ( $N = 2$ ,  $N = 3$ , and  $N = 4$ ). Arham Namazi and Mozaffari (2023): performed 18 experiments with a fixed weir height ( $P = 15$  cm) configured as a single cycle ( $N = 1$ ).

In total, 233 experimental results were collected for soft computing-based modeling. Among these, 175 samples (75%) were used for model training, and 58 samples (25%) were allocated for testing the developed models.

## Results and Discussion

Violin plots for both measured and predicted data inferred by various machine learning models are presented. Violin plots are typically used to compare the distribution of data across different groups in terms of their shape. Additionally, a small box plot is embedded within each violin plot, where the ends of the rectangle represent the first and third quartiles, and the central point denotes the median. It can be observed that all three models—RF, SVM, and MARS—predict similar first and third quartiles and medians, compared to the measured data. In contrast, the first or third quartiles in the equations proposed by Arham Namazi and Mozaffari (2023) and Equation 18 show significant deviations from the measured values. Furthermore, from the perspective of the overall data distribution, the SVM and MARS algorithms demonstrate distributions more similar to the measured data compared to the RF algorithm. This highlights the superior predictive capability of the support vector machine (SVM) approach.

## Conclusions

Labyrinth weirs are consistently proposed as an effective solution for enhancing flood discharge efficiency, particularly in cases where space for weir construction is limited. These weirs have a longer crest length compared to linear weirs, allowing floods to pass at shallower depths. Due to the complex relationship between the discharge coefficient and its associated parameters, empirical equations often fail to predict the discharge coefficient with acceptable accuracy.

In this study, three different machine learning models were developed to predict the discharge coefficient of semicircular labyrinth weirs with harmonic designs. The results confirm the advantages of the Support Vector Machine (SVM) algorithm. Key findings of the study are summarized as follows:

1. **Parameter Sensitivity Analysis:** To minimize prediction errors in the machine learning models, a sensitivity analysis was conducted to identify the relative head importance of different input parameters. Based on this analysis, five input combinations were designed and applied to the machine learning models.
2. **Optimal Input Combination:** Statistical comparisons between predicted and experimental data revealed that the optimal input combination effectively predicted the discharge coefficient for this type of weir.
3. **Model Performance:** Using the best input combination, the results showed that the SVM and MARS algorithms outperformed tree-based models, such as Random Forest (RF), in prediction accuracy for harmonic weirs with varying cycles.
4. **MARS Model Evaluation:** Although the MARS model performed well, comparisons with other regression models from previous studies demonstrated that MARS delivered satisfactory and improved accuracy over those models.

**Keywords:** Machine Learning, SHAP Sensitivity Analysis, Discharge Coefficient, Labyrinth Weirs

