

Research Paper

Performance Assessment of Various Data-Mining Methods to Determine Depth Velocity Profile at Submerged Hydraulic Jump

Mahdi Majedi-Asl^{1*}, Mehdi Fuladipanah²

¹ Associate Professor, Department of Civil Engineering, University of Maragheh, Maragheh, Iran, mehdi.majedi@gmail.com

² Assistant Professor, Department of Civil Engineering, Ramhormoz Branch, Islamic Azad University, Ramhormoz, Iran. fuladipanah@gmail.com



[10.22125/IWE.2022.162560](https://doi.org/10.22125/IWE.2022.162560)

Received:
December,02,2020
Accepted:
May,01,2020
Available online:
December 11, 2022

Keywords:
Flow Velocity,
Gamma test,
Hydraulic jump, Gene
Expression
Programming,
Support Vector
Machine.

Abstract

Vertical velocity distribution at hydraulic jump is one of the challenging and significant issues among researchers because of the complexity of measurement and calculations. In this research, the application of the SVM and GEP intelligent models has been considered to determine the vertical velocity profile at the submerged hydraulic jump downstream of a sluice gate. Laboratory measured data of number 312 has been used in the simulation. Using dimensional analysis, dimensionless input parameters were introduced to models including upstream Froud number (Fr_1), Tail-water Froud number (Fr_3), the ratio of upstream flow depth to the tail-water depth ($\frac{y_1}{y_3}$), the ratio of the gate opening to the channel width ($\frac{w}{b}$), and the ratio of vertical distance from the channel bed to the channel width ($\frac{z}{b}$). Using the gama test, all five parameters were determined as the optimum combination to simulate velocity profile. Of two Nu-SVM and C-SVM classification models, the first one was opted as optimum model of the SVM algorithm with RBF Kernel function with the setting parameters γ and Nu of values 1.2 and 0.486, respectively. The performance of the Nu-SVM and the GEP intelligent, models were assessed using statistical criteria. The results showed that the values of (RMSE, R^2 , $(\frac{u}{U_1})_{DDR(max)}$) indices for the test phase of the Nu-SVM and the GEP algorithms are (0.09588,0.9770,0.4489) and (0.1161,0.9718,0.3588) respectively; illustrating the superiority of the Nu-SVM algorithm. Also, according to the gama test, the arrangement of the effective dimensionless parameters on the velocity profile is $\frac{z}{b}$, Fr_1 , $\frac{w}{b}$, Fr_3 and $\frac{y_1}{y_3}$.

1. Introduction

The control of flow velocity to prevent damage and destruction in hydraulic system and structures is one of the significant aspects of their safety design. As a non-structural type of energy dissipater

* **Corresponding Author:** Mahdi Majedi-Asl

Address: Department of Civil Engineering,
University of Maragheh, Maragheh, Iran

Email: mehdi.majedi@gmail.com

Tel: 09143222205

method, hydraulic jump has more application in this field. Free and submerged states are two types of hydraulic jump that may be observed in hydraulic conduits. A vast number of experimental, empirical and theoretical equations have been developed to model the depth velocity profile of submerged jump. But, acceptable precise have not been gained due to the inability of these equations to quantify some effective parameters on depth velocity distribution. Machine learning algorithms as a type of data mining approach, are capable to extract the complex and hidden relationships between dependent and independent variables. In this research work, the depth profile of the flow velocity in the submerged hydraulic jump formed downstream of the sliding valve using SVM and GEP algorithms based on the laboratory data of Bijankhan et al. (2017) has been done.

1. Materials and Methods

Experimental dataset of number 312 were utilized to simulate the depth velocity profile of submerged hydraulic jump from Bijankhan et al. (2017). All experiments have been done in a plexiglass flume of rectangular section with length of 7 m, width 1.179 m and height 1 m. Support Vector Machine (SVM) and Gene Expressing Programming (GEP) were employed to predict depth velocity profile. Five dimensionless parameters including Fr_1 , Fr_3 , $\frac{y_1}{y_3}$, $\frac{z}{b}$, $\frac{w}{b}$ were extracted using dimensional analysis to predict velocity profile. All five predictors were normalized between (0,1) and a Gamma test was implemented to optimize combination of above-mentioned parameters. The values of Γ demonstrated that all independent parameters have remarkable role to predict velocity profile. The effect order of parameters is $\frac{z}{b}$, Fr_1 , $\frac{w}{b}$, Fr_3 and $\frac{y_1}{y_3}$, respectively, based on the gamma test. Root mean square error (RMSE), determination coefficient (R^2) and normalized Developed Discrepancy Ratio (Z_{DDR}) were used to evaluate models' performance.

2. Results

Nu-SVM and C-SVM classification approaches were assessed to predict velocity profile. The values of RMSE, R^2 and Z_{DDR} indices of superior approach (i.e. Nu-SVM) during training and testing phases were obtained (0.08376, 0.9825, 0.3178) and (0.09588, 0.9770, 0.4489), respectively. The share of training and testing phases of datasets were 80% and 20%, respectively. Of all data, the share of training and testing stages for GEP algorithm were 75% and 25%. The values of RMSE, R^2 and Z_{DDR} indices for GEP model during training and testing phases were calculated as (0.2183, 0.8609, 0.2045) and (0.1161, 0.9718, 0.3588), respectively.

3. Discussion and Conclusion

Ability of machine learning methods including SVM and GEP was evaluated to predict the vertical velocity distribution in submerged hydraulic jump using experimental data gathered from a rectangular section. Although assessment indices proved both models capability, but the outcome of SVM was accurate.

4. Six important references

1. Bijankhan, M., S. Kouchakzadeh and G. Belaud. 2017. Application of the submerged experimental velocity profiles for the sluice gate's stage-discharge relationship. *Flow Measurement and Instrumentation*, 54: 97-108.
2. Gu, S., Bo F., M. Luo, E. Kazemi, Y. Zhang and J. Wei. 2019. SPH Simulation of Hydraulic Jump on Corrugated Riverbeds. *Applied science*, 9(436): 1-18.
3. Houichi, L., N. Dechemi, S. Heddam and B. Achour. 2013. An evaluation of ANN methods for estimating the lengths of hydraulic jumps in U-shaped channel. *Journal of Hydroinformatics*, 15(1): 147-154.

4. Jesudhas, V., V. Roussinova, R. Balachandar and R. Barron. 2017. Submerged Hydraulic Jump Study Using DES. *Journal of hydraulic engineering*, 143(3): 1-11.
5. Naseri, M. and F. Othman. 2012. Determination of the length of hydraulic jumps using artificial neural networks. *Advances in Engineering Software*, 48: 27-31.
6. Noori, R., A. Khakpour, B. Omidvar, and A. Farokhnia. (2010). Comparison of ANN and principal component analysis-multivariate linear regression models for predicting the river flow based on developed discrepancy ratio statistic. *Expert Systems with Applications*, 37: 5856-5862.

Conflict of Interest

Authors declared no conflict of interest.

ارزیابی عملکرد روش‌های داده کاوی مختلف در تعیین پروفیل عمقی سرعت در پرش هیدرولیکی مستغرق

مهدی ماجدی اصل^۱، مهدی فولادی پناه^۲

تاریخ ارسال: ۱۳۹۹/۰۹/۱۲

تاریخ پذیرش: ۱۴۰۰/۰۲/۱۱

مقاله پژوهشی

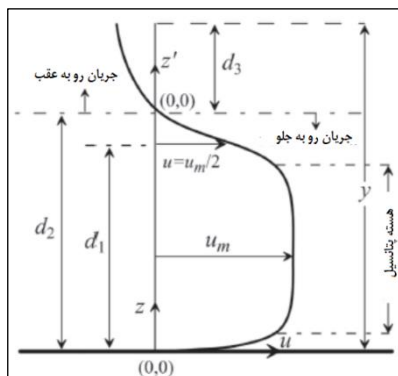
چکیده

توزیع عمقی سرعت در پرش هیدرولیکی به دلیل پیچیدگی‌های اندازه‌گیری و محاسباتی به عنوان یکی از موضوع‌های مهم و در حال بررسی میان پژوهشگران مطرح می‌باشد. در این تحقیق، عملکرد مدل‌های هوشمند GEP و SVM در تعیین توزیع عمقی سرعت در پرش هیدرولیکی مستغرق در پایین دست دریچه کشویی مورد بررسی قرار گرفت. بدین منظور از ۳۱۲ داده آزمایشگاهی استفاده شده است. با استفاده از تحلیل ابعادی پارامترهای بی‌بعد ورودی به مدل‌های هوشمند شامل عدد فرود بالادست (Fr_1)، عدد فرود پایاب (Fr_3)، نسبت عمق بالادست دریچه به عمق پایاب ($\frac{y_1}{y_3}$)، نسبت بازشدگی دریچه به عرض کانال ($\frac{w}{b}$) و نسبت عمق اندازه‌گیری به عرض کانال ($\frac{z}{b}$) معین شدند. با استفاده از آزمون گاما، استفاده از هر پنج پارامتر بی‌بعد به عنوان بهترین ترکیب برای پیش‌بینی توزیع عمقی سرعت تعیین شد. داده‌های آزمایشگاهی اندازه‌گیری شده نرمال‌سازی شدند. از بین دو کلاس طبقه‌بندی Nu-SVM و C-SVM، الگوریتم اول به ازای مقدار پارامترهای تنظیمی γ و Nu به ترتیب برابر با ۱/۲ و ۰/۴۸۶ با تابع کرنل از نوع RBF به عنوان الگوریتم برتر برای SVM انتخاب شد. عملکرد مدل‌های هوشمند Nu-SVM و GEP با استفاده از شاخص‌های آماری مورد ارزیابی قرار گرفت. نتایج نشان داد که مقدار شاخص‌های ($RMSE$, R^2) ($\frac{u}{U_1}$) $_{DDR(max)}$ در فرآیند آزمون برای الگوریتم‌های Nu-SVM و GEP به ترتیب (۰/۴۴۸۹، ۰/۹۷۷۰، ۰/۰۹۵۸۸) و (۰/۳۵۸۸، ۰/۹۷۱۸، ۰/۱۱۶۱) هستند که نشان از برتری الگوریتم Nu-SVM دارد. همچنین طبق آزمون گاما، ترتیب تاثیرگذاری پارامترهای بی‌بعد موثر بر روی پروفیل عمقی سرعت به صورت $\frac{z}{b}$ ، Fr_1 ، $\frac{w}{b}$ ، Fr_3 و $\frac{y_1}{y_3}$ به دست آمد.

واژه‌های کلیدی: آزمون گاما، الگوریتم برنامه‌ریزی بیان ژن، الگوریتم ماشین بردار پشتیبان، پرش هیدرولیکی، سرعت جریان.

^۱ دانشیار، گروه مهندسی عمران، دانشگاه مراغه، مراغه، ایران (نویسنده مسئول). mehdi.majedi@gmail.com

^۲ استادیار، گروه عمران، واحد رامهرمز، دانشگاه آزاد اسلامی، رامهرمز، ایران. fuladipanah@gmail.com



شکل (۲): بیان نمادین از پروفیل عمقی سرعت در پرش هیدرولیکی مستغرق (Rajaratnam and Subramanya, 1967)

رابطه‌های ارائه شده توسط Rajaratnam and Subramanya (1967) برای ناحیه‌های رو به جلو و رو به عقب به ترتیب در معادله‌های (۱) و (۲) بیان شده‌اند:

$$\frac{u}{u_m} = \begin{cases} 1 & 0 \leq \frac{z}{d_1} \leq 0.66 \\ 0.6375 \left(\frac{z}{d_1}\right)^2 - 2.5972 \left(\frac{z}{d_1}\right) + 2.45 & 0.66 < \frac{z}{d_1} \leq 1.486 \end{cases} \quad (1)$$

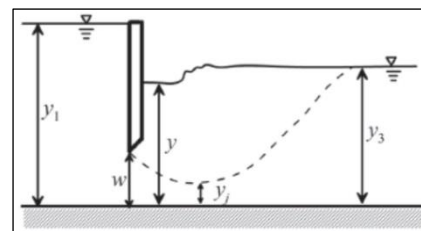
$$u = \frac{0.5q}{C_c W} \left[1 - 2 \frac{z - C_c W}{y - C_c W} \right] \quad (2)$$

که در آن u سرعت جریان در عمق Z بر حسب متر بر ثانیه، q دبی در واحد عرض بر حسب متر مکعب بر ثانیه بر متر، w ارتفاع بازشدگی دریاچه بر حسب متر، C_c ضریب انقباض دریاچه، Z ارتفاع از کف مجرا در راستای عمود بر حسب متر، u_m سرعت بیشینه در هسته پتانسیل بر حسب متر بر ثانیه، y عمق جریان بر حسب متر، d_1 عمق متناظر با سرعت $\frac{u_m}{2}$ و d_3 عمق متناظر با سرعت صفر هستند.

پژوهش‌های مختلفی بر روی پیش‌بینی ویژگی‌های پرش هیدرولیکی توسط پژوهشگران انجام شده است. عباسپور و همکاران (۱۳۸۸) پژوهش آزمایشگاهی در خصوص تاثیر بستر موج‌دار روی ویژگی‌های پرش هیدرولیکی از جمله پروفیل سرعت انجام دادند. بررسی پروفیل‌های سرعت اندازه‌گیری شده در مقاطع مختلف پرش هیدرولیکی نشان داد که این پروفیل‌ها مشابه بوده و با پروفیل جت آب بر روی بستر صاف متفاوت می‌باشند. غزالی و همکاران (۱۳۹۱) در یک پژوهش آزمایشگاهی به بررسی و مطالعه پروفیل سرعت در پرش هیدرولیکی روی بستر با زبری مثلی پرداختند و به این نتیجه رسیدند که شکل زبری روی پروفیل عمقی سرعت موثر است. هنر و پورحمزه

مقدمه

کنترل سرعت جریان آب در سیستم‌ها و سازه‌های هیدرولیکی به منظور جلوگیری از بروز خسارت و تخریب، یکی از مهم‌ترین جنبه‌های طراحی آنها به شمار می‌رود. پرش هیدرولیکی به عنوان یکی از گزینه‌های غیرسازه‌ای برای کاهش انرژی جریان کاربرد بسیار زیادی در این زمینه دارد. پرش هیدرولیکی جریان متغیر سریعی است که جریان ضمن گذر از وضعیت فوق‌بحرانی به زیربحرانی مقدار قابل توجهی از انرژی جنبشی خود را از به دلیل کاهش سرعت از دست می‌دهد. پرش هیدرولیکی در دو حالت آزاد و مستغرق رخ می‌دهد. در نوع آزاد، عمق پایاب برابر با عمق ثانویه پرش هیدرولیکی و در حالت مستغرق، عمق پایاب بیشتر از عمق ثانویه مربوط به پرش آزاد است. نمونه‌ای از پرش هیدرولیکی مستغرق در شکل (۱) نشان داده شده است. Rajaratnam and Subramanya (1967) در مقطع انقباض، پروفیل عمقی سرعت را به صورت نمادین در قالب شکل (۲) نشان داده است. در قسمتی از جریان به عمق d_2 ، جریان رو جلو شکل می‌گیرد. در عمقی معادل با d_3 ، جریان رو عقب شکل می‌گیرد که دارای سرعت منفی است. در داخل محدوده‌ی جریان رو به جلو، ناحیه‌ی پتانسیل قرار دارد که همانند سیال ایده‌آل، پروفیل عمقی سرعت در آن یکنواخت است.



شکل (۱): میدان جریان در پرش هیدرولیکی مستغرق (Rajaratnam and Subramanya, 1967)



نشان داد مدل برتر، مقادیر آزمایشگاهی را با دقت قابل قبولی پیش‌بینی می‌کند. هوشیاری پور و همکاران (۱۳۹۸) با استفاده از مدل‌سازی عددی تاثیر زاویه واگرایی دیواره حوضچه‌ی آرامش و موقعیت آب پایه را روی خصوصیات پرش هیدرولیکی بررسی کردند. از کاربردهای الگوریتم‌های هوشمند می‌توان به پژوهش‌های انجام شده توسط فرزین و ولیخان (۱۳۹۹)، نیک‌پور و محمودی‌بابان (۱۳۹۷)، ماجدی‌اصل و فولادی‌پناه (۱۳۹۷)، فولادی‌پناه و همکاران (۱۳۹۹) اشاره نمود. (Naseri and Othman (2012) با استفاده از شبکه عصبی مصنوعی (ANN) طول پرش هیدرولیکی را تخمین زدند و با مقایسه نتایج آن با روابط تجربی دریافتند که ANN منجر به تخمین بهتری از طول پرش هیدرولیکی می‌شود. طول پرش هیدرولیکی با استفاده از مدل‌های داده‌مبنای شبکه عصبی مصنوعی عمومی (GRNN) و شبکه عصبی پرسپتون چند لایه (MLPNN) توسط Houichi et al. (2013) شبیه‌سازی شد. نتایج حاصل از این پژوهش، قابلیت بیشتر مدل عصبی MLPNN در پیش‌بینی طول پرش هیدرولیکی بود. Jesudhas et al. (2017) با استفاده از شبیه‌سازی گردابه‌ای جدا شده (DES) پروفیل عمقی سرعت همراه با میزان افت انرژی جریان را برای پرش هیدرولیکی مستغرق با عدد فرود $8/2$ شبیه‌سازی کردند. نتایج آنها نشان داد DES قابلیت شبیه‌سازی سه بعدی میدان جریان را با دقت بسیار زیاد دارا می‌باشد. در پژوهش دیگری، Gu et al. (2019) با استفاده از مدل عددی هیدرودینامی ذره هموار (SPH)، میدان سرعت، افت انرژی و جبهه موج پرش هیدرولیکی را روی بستر موجدار شبیه‌سازی نمودند. نتایج پژوهش آنها دال بر انطباق بسیار مناسب نتایج مدل SPH با مقادیر آزمایشگاهی بود. مدل SVM توسط Saghebian and Dragomir-Stanciu (2019) برای پیش‌بینی افت انرژی در پرش هیدرولیکی در بستر زبر مورد استفاده قرار گرفت. نتایج آنها نشان داد این مدل توانایی بسیار مناسبی در شبیه‌سازی افت انرژی در پرش هیدرولیکی دارد. معادله‌های تجربی، نیمه‌تجربی و تئوریکی مختلفی برای مدلسازی ویژگی‌های پرش هیدرولیکی مستغرق از جمله پروفیل عمقی سرعت ارائه شده‌اند؛ ولی به دلیل عدم توانایی معادلات در کمی نمودن

(۱۳۹۱) با استفاده از شبکه عصبی مشخصات پرش هیدرولیکی را در حوضچه‌ی آرامش با دیواره‌ی همگرا و شیب کف منفی پیش‌بینی نمودند. مقادیر بالای به دست آمده برای ضریب رگرسیون، بیانگر همبستگی نزدیک بین مقادیر خروجی مدل شبکه عصبی با داده‌های آزمایشگاهی بود. عباسپور (۱۳۹۳) به تخمین مشخصات پرش هیدرولیکی بر روی بستر زبر با استفاده از شبکه عصبی مصنوعی و برنامه‌ریزی ژنتیک پرداخت. وی با تحلیل عمق و طول غلتان پرش بر روی بسترهای زبر، مدل شبکه عصبی و برنامه‌ریزی ژنتیک پرش را شبیه‌سازی نمود و نتایج آن را با روابط تجربی مقایسه کرد. کمتر بودن خطای محاسباتی مدل‌های عصبی و برنامه‌ریزی ژنتیک نسبت به روابط تجربی از نتایج این پژوهش بود. پورعبدالله و همکاران (۱۳۹۴) تاثیر بستر زبر و شیب معکوس را روی پروفیل آب و سرعت در پرش هیدرولیکی با عدد فرود بین $4/9$ تا $7/8$ به صورت آزمایشگاهی مطالعه نمودند. نتایج حاکی از تشابه پروفیل‌های سرعت اندازه‌گیری شده و تفاوت آنها با پروفیل جت آب بر روی بستر صاف بود. روشنگر و ولی‌زاده (۱۳۹۶) در پژوهش خود به پیش‌بینی خصوصیات پرش هیدرولیکی (عمق متناوب و افت انرژی) با استفاده از SVM پرداختند و بیان داشتند الگوریتم SVM نسبت به معادلات نیمه‌تجربی دقت و عملکرد بسیار مطلوب‌تری برای پیش‌بینی ویژگی‌های پرش هیدرولیکی دارد. تبریزی و ثاقبیان (۱۳۹۸) با استفاده از الگوریتم SVM به بررسی و پیش‌بینی ویژگی‌های پرش هیدرولیکی (عمق و طول پرش) در کانال‌های با مقاطع مرکب با بستر زبر پرداختند و نشان دادند الگوریتم SVM قابلیت مناسبی در این زمینه دارد. پورعبدالله و همکاران (۱۳۹۹) تاثیر شیب منفی، زبری بستر و پله انتهایی را روی خصوصیات پرش هیدرولیکی از جمله پروفیل سرعت به صورت آزمایشگاهی مورد بررسی قرار دادند. نتایج پژوهش آنها نشان داد سه عامل شیب منفی، زبری بستر و پله مثبت انتهایی باعث کاهش انحراف معیار استاندارد نوسانات فشار نسبت به حالت کلاسیک می‌شود. علیزاده و همکاران (۱۳۹۸) با استفاده از سیستم استنباط فازی عصبی تطبیقی با الگوریتم ژنتیک (ANFIS-GA) طول پرش هیدرولیکی را روی بسترهای زبر شیب‌دار مدلسازی کردند. نتایج آنها

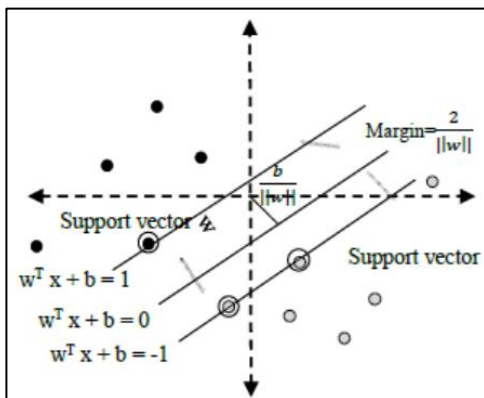
جدول (۱): محدوده متغیرهای آزمایشگاهی در پژوهش حاضر

(Bijankhan et al., 2017)

ردیف	متغیر	نماد	یکا	محدوده
۱	دبی	Q	Lit/s	۳۶۹/۵ تا ۱۲/۰۴
۲	فاصله از کف فلوم	Z	mm	۰ تا ۲۰۹/۵
۳	سرعت	u	cm/s	۱۷۰/۹۸ تا ۳۹/۶۲
۴	عمق آب قبل از دریاچه	y_1	mm	۱۹۵/۱ تا ۹۲
۵	عمق آب بعد از پرش	y_3	mm	۱۷۷/۴ تا ۸۴/۵
۶	فاصله افقی از دریاچه	X	m	۰ تا ۵

الگوریتم ماشین بردار پشتیبان

الگوریتم SVM یکی از مدل‌های هوشمند یادگیری با نظارت است که توسط Vapnik ارائه شده است (Vapnik, 1995). مبنای این الگوریتم، استفاده از فرآیندهای آموزش و آزمون است ولی برخلاف دیگر روش‌های هوشمند، به جای کاهش خطای محاسبه‌ای ریسک عملیاتی عدم تقسیم‌بندی درست را به عنوان تابع هدف قرار می‌دهد و مقدار بهینه‌ی آن را به دست می‌آورد. این کار بر مبنای دسته‌بندی خطی داده‌ها انجام می‌شود. خط مقسم بین داده‌ها طوری انتخاب می‌شود که بیشترین حاشیه‌ی اطمینان بین دسته‌ها ایجاد شود. این فرایند در شکل (۴) به طور نمادین نشان داده شده است. در الگوریتم SVM، هدف یافتن خطی است که از داده‌های موجود در دو کلاس بیشترین فاصله یا کمترین ریسک را داشته باشد. سپس دو صفحه‌ی مرزی موازی با صفحه‌ی مقسم به گونه‌ای رسم می‌شود که بیشترین فاصله را از صفحه‌های تقسیم‌کننده داشته باشد.

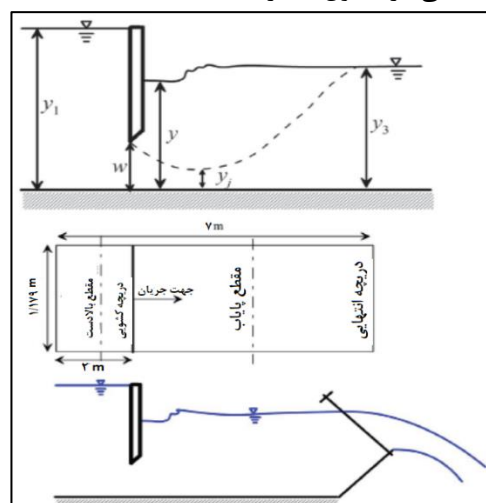


شکل (۴): بردارهای پشتیبان و دسته‌بندی داده‌ها در SVM

برخی پارامترهای موثر در پیش‌بینی خصوصیات پروفیل عمقی سرعت در پرش هیدرولیکی مستغرق، از دقت زیادی برخوردار نیستند. در این وضعیت، استفاده از الگوریتم‌های هوشمند که قادر به استخراج روابط پیچیده و مخفی بین متغیرهای مستقل و وابسته هستند و کاربرد آنها توسط پژوهشگرهای مختلف به اثبات رسیده است می‌تواند دقت مدلسازی را به طور قابل توجهی افزایش دهد. در این پژوهش، پروفیل عمقی سرعت جریان در پرش هیدرولیکی مستغرق شکل گرفته در پایین‌دست دریاچه کشویی با استفاده از الگوریتم‌های SVM و GEP بر مبنای داده‌های آزمایشگاهی (Bijankhan et al., 2017) انجام شده است.

مواد و روش‌ها

داده‌های آزمایشگاهی پروفیل عمقی سرعت در این پژوهش از کار پژوهشی (Bijankhan et al., 2017) جمع‌آوری شده است. در شکل (۳) به صورت نمادین پلان و دید جانبی سرریز همراه با پرش هیدرولیکی نشان داده شده است. آنها آزمایش خود را در یک فلوم از جنس پلاکسی-گلاس به طول ۷ متر، عرض ۱/۱۷۹ متر و ارتفاع ۱/۰ متر انجام دادند. عمق بازشدگی دریاچه در آزمایش‌ها ۴۶/۵ و ۶۸/۳ میلیمتر لحاظ شدند. مقدار دبی جریان توسط دبی-سنج با دقت $\pm 5\%$ اندازه‌گیری شده است. دقت عمق‌سنج مورد استفاده ۰/۱ میلیمتر می‌باشد. محدوده‌ی متغیرهای آزمایشگاهی در جدول (۱) ارائه شده است.



شکل (۳): دید بالا و نمای جانبی از سرریز و پرش هیدرولیکی مستغرق

(Bijankhan et al., 2017)

نوع تابع کرنل بر مبنای آزمون و خطا انجام می‌شود (Vapnik, 1995).

جدول (۲): انواع تابع کرنل

نام تابع کرنل	تابع
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	$K(x_i, x_j) = \exp[-\frac{\ x_i - x_j\ ^2}{2\sigma^2}]$
تابع پایه شعاعی نمایی	$K(x_i, x_j) = \tanh[-\alpha(x_i, x_j) + c]$

الگوریتم برنامه‌ریزی بیان ژن^۱

الگوریتم گردشی GEP در گروه الگوریتم‌های مبتنی بر نظریه‌ی تکامل داروین است که در آن تعریف تابع هدف بر مبنای معیارهای کیفی انجام می‌شود. تابع هدف به منظور مقایسه‌ی جواب‌های مختلف مسئله در قالب فرآیندی گام به گام تصحیح ساختار داده‌ها برای دستیابی به جواب بهینه به کار گرفته می‌شود. بدین منظور، ضمن کدگذاری افراد به صورت رشته‌های خطی با طول ثابت (ژنوم یا کروموزوم)، آنها به فرم نهادهای غیرخطی با اندازه‌ها و شکل‌های مختلف (بیان درختی) نمایش داده می‌شوند. تولید جمعیت اولیه از راه‌حل‌ها به کمک فرآیند کاملاً تصادفی یا بر مبنای مقداری از اطلاعات موجود از مسئله اولین مرحله‌ی این الگوریتم هوشمند است. با به کارگیری یک تابع برازش مناسب، کروموزوم‌ها در قالب بیان درختی نشان داده می‌شوند. در صورت دستیابی به کیفیت رضایت‌بخش از یک راه‌حل بر مبنای تابع برازش یا رسیدن نسل‌ها به تعداد معین، تکامل متوقف می‌شود. در غیر این صورت، ضمن حفظ نسل حاضر، نخبه‌گزینی انجام می‌شود. براساس گزینش، بهترین افراد شانس بیشتری برای تولید فرزندان دارند. انتظار از این روند با تکرار برای چند نسل و با پیشروی نسل به جلو، بهبود کیفیت جمعیت است. مراحل پنجگانه‌ی انجام الگوریتم GEP به قرار زیر است: (۱) تعیین تابع برازش؛ (۲) انتخاب مجموعه ترمینال‌ها و مجموعه توابع برای ایجاد کروموزوم‌ها؛ (۳) انتخاب ساختار کروموزوم‌ها؛ (۴) انتخاب تابع پیوند؛ (۵) انتخاب عملگرهای ژنتیکی و نرخ آنها. فلوجارت الگوریتم

کمترین فاصله از صفحه تا نمونه دارای برچسب +۱ معادل با کمترین فاصله از آن صفحه با برچسب -۱ است. در واقع حاشیه با دو برابر کردن این فاصله محاسبه می‌شود. معادله‌ی این صفحه به قرار زیر است:

$$\mathbf{W} \cdot \mathbf{X} + b = 0 \quad (1)$$

که در این معادله $\mathbf{W} = \{w_1, w_2, \dots, w_n\}$ برداری است که شمار مولفه‌های آن معادل با شمار ویژگی‌هاست و میزان آن ثابت است. در فضای دو بعدی که داده‌ها با دو ویژگی و یک برچسب کلاس تعریف می‌شوند معادله‌ی (۱) همانند معادله‌ی (۲) با فرض $w_0 = b$ بازنویسی می‌شود:

$$w_0 + w_1 x_1 + w_2 x_2 = 0 \quad (2)$$

بر این مینا، نمونه‌ها (نقطه‌ها) واقع در بالای صفحه در نامعادله‌ی (۳) و نقطه‌های زیر این صفحه در نامعادله‌ی (۴) صدق می‌کنند:

$$w_0 + w_1 x_1 + w_2 x_2 > 0 \quad (3)$$

$$w_0 + w_1 x_1 + w_2 x_2 < 0 \quad (4)$$

با تنظیم \mathbf{W} و b می‌توان نوشت:

$$w_0 + w_1 x_1 + w_2 x_2 \geq +1 \text{ if } y_i = +1 \quad (5)$$

$$w_0 + w_1 x_1 + w_2 x_2 \leq -1 \text{ if } y_i = -1 \quad (6)$$

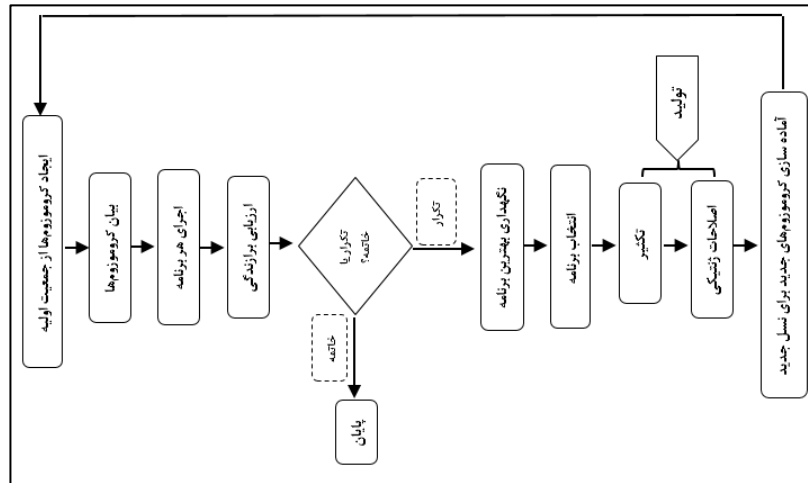
این بدین معناست که هر نمونه‌ی واقع شده روی صفحه‌ی H_1 متعلق به کلاس +۱ و هر نمونه واقع شده زیر صفحه‌ی H_2 متعلق به کلاس -۱ است. همه‌ی نمونه‌هایی که به درستی و با دقت روی صفحه‌های H_1 و H_2 قرار دارند بردار پشتیبان نامیده می‌شوند. یک روش معمول برای حل مسئله‌های غیرخطی استفاده از تابع‌های کرنل است. این تابع‌ها به صورت حاصل ضرب داخلی داده‌ها تعریف می‌شوند. در حقیقت، با انتقال غیرخطی از فضای داخلی به فضای ویژگی‌ها با ابعاد بزرگ‌تر، می‌توان مسئله‌ها را به صورت خطی جداسازی کرد. با انتقال نمونه‌ها از فضای داخلی به فضای ویژگی‌ها، جداکننده‌ی غیرخطی به صورت خطی تبدیل می‌شود (معادله‌ی (۷)):

$$\mathbf{W}^T \Phi(x_j) + b = 0 \quad (7)$$

که در آن \mathbf{W}^T ترانهاده‌ی بردار \mathbf{W} و $\Phi(x_j)$ تابع کرنل هستند که انواع آن در جدول (۲) ارائه شده است. انتخاب

^۱ Gene Expression Programming (GEP)

GEP در شکل (۵) نشان داده شده است (فولادی پناه و همکاران، ۱۳۹۹).



شکل (۵): روندنمای الگوریتم GEP

محدود است و تقریباً برابر با ۱۰ در نظر گرفته می‌شود. در این صورت الگوریتم آزمون گاما مقدار زیر را محاسبه می‌کند:

$$\delta_M(k) = \frac{1}{M} \sum_{i=1}^M |x_{N[i,k]} - x_i|^2 \quad (9)$$

که در این معادله نماد $| |$ بیانگر فاصله‌ی اقلیدسی است و در تابع

$$\gamma_M(k) = \frac{1}{2M} \sum_{i=1}^M (y_{N[i,k]} - y_i)^2 \quad (10)$$

عبارت $y_{N[i,k]}$ نشان دهنده‌ی مقدار خروجی متناظر با مقدار $x_{N[i,k]}$ است. برای محاسبه‌ی مقدار آماره‌ی Γ ، کافی است یک خط رگرسیونی با معادله‌ی

$$\gamma_M(k) = \Gamma + A\delta_M(k) \quad (11)$$

از نقاط $(\delta_M(k), \gamma_M(k))$ به ازای $1 \leq k \leq p$ برازش داده شود. مقدار Γ در واقع عرض از مبدا خط برازش یافته‌ی معادله‌ی (۱۱) خواهد بود. پارامتر A که نشان‌دهنده‌ی شیب خط برازش معادله‌ی (۱۱) است بیانگر پیچیدگی مدل رگرسیونی برازش داده شده روی داده‌ها است. حال اگر فرض شود N متغیر ورودی بر روی پدیده‌ای موثر باشند می‌توان حالت 2^{N-1} برای ساخت مدل بین ورودی-خروجی به دست آورد که یافتن مدل برتر و بهینه بسیار وقت‌گیر خواهد بود. آزمون Γ می‌تواند برای یافتن مدل برتر مورد استفاده قرار گیرد. مراحل کار به صورت گام به گام خواهد بود. ابتدا به ازای تمام پارامترها مقدار Γ به دست می‌آید. سپس یکی از

آزمون گاما^۱

فرآیند GT ابزاری برای مدل‌سازی غیرخطی برای تشخیص موثرترین متغیرها در مدل‌سازی و تعیین بهترین ترکیب از متغیرهای ورودی به مدل است (شریفی گرم‌دره و همکاران، ۱۳۹۸). در این فرآیند با استفاده از محاسبه‌ی آماره‌ی گاما (Γ)، برآوردی از بهترین مقدار میانگین مربعات خطا روی خروجی مدل انجام می‌شود. اگر x به عنوان برداری از فضای R^m و مقدار y به عنوان خروجی از فضای R فرض شوند و همچنین فرض بر این باشد که بردارهای x قابلیت پیشگویی فاکتورهای مورد استفاده‌ی موثر در خروجی y را داشته باشند در این صورت مطابق تنها فرض باقیمانده تحت روابط سیستم کنترل، رابطه‌ی زیر دنبال خواهد شد:

$$y = f(x_1, \dots, x_m) + \Gamma \quad (8)$$

که در این معادله f یک تابع نرم و Γ یک متغیر تصادفی است که نشان‌دهنده‌ی مقدارهای اغتشاش است. آماره‌ی Γ از قسمتی از واریانس خروجی که نمی‌تواند برای یک مدل نرم محاسبه شود به دست می‌آید. فرض می‌شود $x_{N[i,k]}$ در بازه‌ی $1 \leq i \leq M$ به ازای $1 \leq k \leq p$ بیانگر k مین همسایه‌ی نزدیک در یک فاصله‌ی اقلیدسی باشد. مقدار p ثابت و

^۱ Gamma Test (GT)

$$X_{\text{normal}} = \frac{X_i - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} \quad (14)$$

که در این معادله، X_{min} و X_{max} به ترتیب نشان دهنده‌ی مقادیر کمینه و بیشینه‌ی متغیر X_i و X_{normal} مقدار متناظر نرمال شده‌ی متغیر X_i است. تعداد داده‌های مورد استفاده در این پژوهش ۳۱۲ داده‌ی آزمایشگاهی است.

تابع‌های ارزیابی عملکرد

سه شاخص مجذور میانگین مربعات خطا (RMSE)، ضریب تبیین (R^2) و آماره‌ی نسبت تفاوت توسعه داده شده^۱ (DDR) برای ارزیابی عملکرد مدل‌های هوشمند به شرح زیر مورد استفاده قرار گرفتند:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N (x_{oi} - x_{pi})^2}{N}} \quad (15)$$

$$R^2 = \frac{\sum_{i=1}^N (x_o^i - \bar{x}_o)(x_p^i - \bar{x}_p)}{\sqrt{[\sum_{i=1}^N (x_o^i - \bar{x}_o)^2][\sum_{i=1}^N (x_p^i - \bar{x}_p)^2]}} \quad (16)$$

$$\text{DDR} = \frac{x_p}{x_o} - 1 \quad (17)$$

که در این روابط x_o و x_p به ترتیب داده‌های مشاهده‌ای و برآورد شده، \bar{x}_o و \bar{x}_p به ترتیب مقادیر متوسط داده‌های مشاهده‌ای و پیش‌بینی شده هستند. با توجه به اینکه شاخص‌های معادله‌های (۱۵) و (۱۶) بیانگر متوسط خطا هستند و ارزیابی در مورد توزیع خطاها ارائه نمی‌دهند برای رفع این محدودیت، (Noori et al. 2010) استفاده از آماره-ی DDR را پیشنهاد دادند. برای مشاهده توزیع خطاها و قضاوت بهتر در خصوص عملکرد مدل پیش‌بینی کننده، محاسبه‌ی تابع گوسی مقادیر DDR به صورت رسم توزیع نرمال استاندارد انجام می‌شود. تمایل بیشتر به سمت محور عمودی مرکزی همراه با مقادیر بزرگتر روی آن محور بیانگر دقت زیاد عملکرد مدل مورد نظر خواهد بود. برای رسم منحنی باید مقادیر استاندارد شده DDR، Z_{DDR} را با استفاده از تابع گوسی مقادیر نرمال شده سرعت عمقی جریان، $(\frac{u}{U_1})_{\text{DDR}}$ ، محاسبه نمود. هر چه منحنی ترسیم شده تمایل بیشتری به محور عمودی داشته باشد و همچنین دارای $(\frac{u}{U_1})_{\text{DDR}}$ بیشتری باشد حاکی از برتری آن مدل پیش‌بینی کننده خواهد بود.

پارامترها حذف می‌شود و مقدار جدید Γ محاسبه می‌شود. در مرحله‌ی بعد، با برگرداندن پارامتر حذف شده به مجموعه پارامترها، پارامتر دیگری حذف می‌شود و مقدار آماره‌ی Γ محاسبه می‌شود. این فرایند به همین منوال ادامه پیدا می‌کند تا بهترین ترکیب از پارامترها با بیشترین مقدار آماره‌ی Γ مشخص شود (Noori et al. 2011). در این پژوهش GT با استفاده از نرم‌افزار WinGamma انجام شده است.

تحلیل ابعادی

ویژگی‌های پروفیل عمقی سرعت تابعی از خصوصیات سیال، خصوصیات مجرای جریان و شرایط هیدرولیکی است. این متغیرها شامل سرعت متوسط جریان در بالادست پرش مستغرق، U_1 ، عمق جریان در بالادست پرش، y_1 ، سرعت متوسط جریان بعد از پرش هیدرولیکی مستغرق، U_3 ، عمق جریان در پایین دست پرش هیدرولیکی مستغرق، y_3 ، لزجت سیال، μ ، چگالی سیال، ρ ، شتاب ثقل، g ، بازشدگی دریچه، w ، عرض مجرای کانال، b ، و عمق آب، z ، است. به عبارتی می‌توان نوشت:

$$u = f(U_1, y_1, U_3, y_3, \mu, g, b, \rho, w) \quad (12)$$

با استفاده از تئوری پی-باکینگهام و فرض g ، ρ و y_1 به عنوان متغیرهای تکراری، معادله‌ی (۱۲) را می‌توان به فرم زیر نوشت:

$$\frac{u}{U_1} = F(\text{Fr}_1, \text{Fr}_3, \frac{y_1}{y_3}, \frac{z}{b}, \frac{w}{b}, \text{Re}) \quad (13)$$

با توجه به مقدار سرعت‌های اندازه‌گیری شده، عدد رینولدز در محدوده‌ی جریان آشفته قرار دارد و به همین دلیل می‌توان از اثرات لزجت و عدد رینولدز صرف‌نظر نمود. بنابراین می‌توان نوشت:

$$\frac{u}{U_1} = f(\text{Fr}_1, \text{Fr}_3, \frac{y_1}{y_3}, \frac{z}{b}, \frac{w}{b}) \quad (12)$$

نرمال‌سازی داده‌ها

محدوده‌ی پارامترهای هیدرولیکی یا هندسی مورد استفاده در این پژوهش متفاوت هستند. به منظور یکسان‌سازی بازه‌ی تغییر هر یک از داده‌ها، از معادله‌ی زیر برای نرمال‌سازی آنها در محدوده‌ی بین ۰ و ۱ استفاده شد:

^۱ Developed Discrepancy Ratio

نتایج و بحث

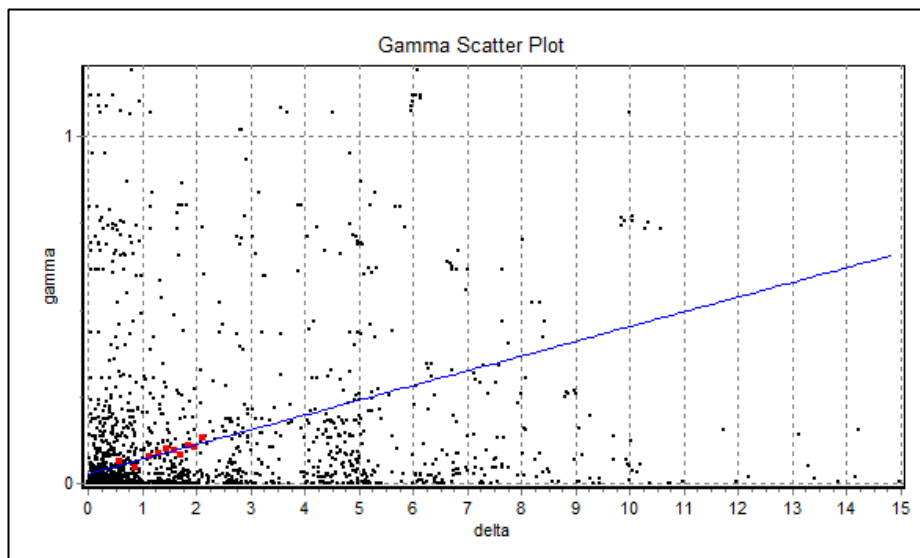
در جدول (۳) نتیجه‌ی مربوط به آزمون Γ به منظور انتخاب بهترین ترکیب ارائه شده است. پارامترهای مورد استفاده در این جدول شامل پارامترهای بی‌بعد ارائه شده در سمت راست معادله‌ی (۱۲) هستند. همانطور که گفته شد برای انجام این کار، مقدار آماره‌ی Γ با اعمال تمام پارامترها محاسبه می‌شود (ترکیب شماره‌ی ۱ جدول (۳)). در مرحله‌ی بعد، یکی از پارامترهای مستقل حذف می‌شود و مقدار آماره‌ی Γ برای آن محاسبه می‌شود (ترکیب‌های شماره‌ی ۲ تا ۵ جدول (۳)).

جدول (۳): نتیجه‌ی آزمون Γ در انتخاب بهترین ترکیب برای

پیش‌بینی $\frac{u}{u_{max}}$ در پرش هیدرولیکی مستغرق

شماره‌ی ترکیب	متغیر	مقدار Γ
۱	All	۰/۰۲۷۶۲
۲	All-Fr ₁	۰/۰۳۵۵۲
۳	All-Fr ₃	۰/۰۳۳۹۳
۴	All- $\frac{y_1}{y_3}$	۰/۰۳۱۴۲
۵	All- $\frac{w}{b}$	۰/۰۳۵۳۲
۶	All- $\frac{z}{b}$	۰/۰۴۶۴۷

ترکیب‌هایی که مقدار آماره‌ی Γ آنها نسبت به ترکیب شماره‌ی ۱ بیشتر باشند به عنوان بهترین ترکیب برای استفاده در مدل‌های پیش‌بینی استفاده خواهند شد. طبق محاسبه‌ی انجام شده در جدول (۳)، مقدار آماره‌ی Γ برای ترکیب ۱ برابر با ۰/۰۲۷۶۲ می‌باشد. حذف پارامترهای Fr₁، Fr₃، $\frac{y_1}{y_3}$ و $\frac{z}{b}$ در ترکیب‌های ۲ تا ۶ به ترتیب منجر به آماره‌ی Γ با مقدارهای ۰/۰۳۵۵۲، ۰/۰۳۳۹۳، ۰/۰۳۱۴۲، ۰/۰۳۵۳۲ و ۰/۰۴۶۴۷ شده است. همانطور که مشاهده می‌شود برای تمام ترکیب‌ها، مقدار آماره‌ی Γ بیشتر از مقدار آماره‌ی Γ ترکیب شماره‌ی ۱ هستند. به عبارتی دیگر، برای پیش‌بینی توزیع عمقی سرعت وجود تمام پارامترهای لازم و ضروری است. با توجه به تغییرات مقدار آماره‌ی Γ ، ترتیب تاثیرگذاری پارامترهای بی‌بعد روی پروفیل عمقی سرعت به صورت $\frac{z}{b}$ ، Fr₁، $\frac{w}{b}$ ، Fr₃ و $\frac{y_1}{y_3}$ است. مقدار آماره‌ی Γ بیشتر به معنی اثرگذارتر بودن آن پارامتر است. همچنین به عنوان نمونه، خط برازش یافته برای ترکیب شماره‌ی ۱ به منظور تعیین مقدار آماره‌ی Γ (عرض از مبدا خط برازش) در شکل (۶) نشان داده شده است.



شکل (۶): عرض از مبدا خط برازش یافته برای ترکیب شماره‌ی ۱ از پارامترهای بی‌بعد ($\Gamma=0/02762$)

همزمان با افزایش حاشیه‌ی جداسازی کلاس‌ها، به دلیل افزایش مقدار Nu ، خطاها افزایش پیدا می‌کند (Fan et al., 2005). در جدول (۴) مقایسه‌ی بین سنجه‌های ارزیابی عملکرد در بهترین ترکیب برای دو روش طبقه‌بندی Nu -SVM و C -SVM ارائه شده است. همانطور که مشاهده می‌شود استفاده از روش طبقه‌بندی Nu -SVM منجر به مقادیر سنجه‌های ارزیابی عملکرد بهتری شده است و به همین دلیل در این پژوهش از این نوع روش طبقه‌بندی استفاده شده است.

جدول (۴): مقایسه بین روش‌های طبقه‌بندی الگوریتم SVM

الگوریتم	فاز	R^2	RMSE	درصد داده‌های مورد استفاده	پارامترهای تنظیمی الگوریتم
Nu-SVM	آموزش	۰/۹۸۲۵	۰/۰۸۳۷۶	۸۰	$\gamma=1$
	آزمون	۰/۹۷۷۰	۰/۰۹۵۸۸	۲۰	$Nu=0/486$ Kernel function: RBF
C-SVM	آموزش	۰/۹۶۲۵	۰/۱۰۵۱	۷۵	$C=98$
	آزمون	۰/۹۷۴۳	۰/۰۹۶۷	۲۵	$\gamma=50$ Kernel function: RBF

به طوری که مقدار بیشینه شاخص $(\frac{u}{u_1})_{DDR}$ در فرآیند آموزش و آزمون به ترتیب ۰/۳۱۷۸ و ۰/۴۴۸۹ است. همچنین تمایل نمودار فرآیند آزمون حول محور قائم بیشتر از فرآیند آموزش است که این واقعیت نیز دلیل بر برتری کارکرد الگوریتم در فرآیند آزمون دارد. مقدار سرعت‌های مشاهده‌ای و محاسبه‌ای توسط الگوریتم Nu -SVM در شکل (۸) نشان داده شده است. همچنین پراکندگی نقاط سرعت‌های مشاهده‌ای و محاسبه‌ای حول خط ۱:۱ در شکل (۹) نشان داده شده است. پراکنش نقاط فرآیند آزمون الگوریتم Nu -SVM حول خط ۱:۱ بهتر از فرآیند آموزش است که موید عملکرد بهتر مدل در فرآیند آزمون است.

الگوریتم SVM
در اجرای الگوریتم SVM، می‌توان از دو طبقه‌بندی C -SVM و Nu -SVM برای دستیابی به تابع خطای کمینه استفاده نمود. در روش کلاس‌بندی C -SVM، فراسنجه‌ی C بر مبنای اطلاع از نویز موجود در داده‌ها انتخاب می‌شود که مقدار آن با روش‌های بهینه‌سازی قابل تعیین است. در روش طبقه‌بندی Nu -SVM، فراسنجه‌ی Nu به عنوان مرز بالایی از خطاها و حد پایین برای بردارهای پشتیبان عمل می‌کند. در نتیجه، با افزایش مقدار Nu

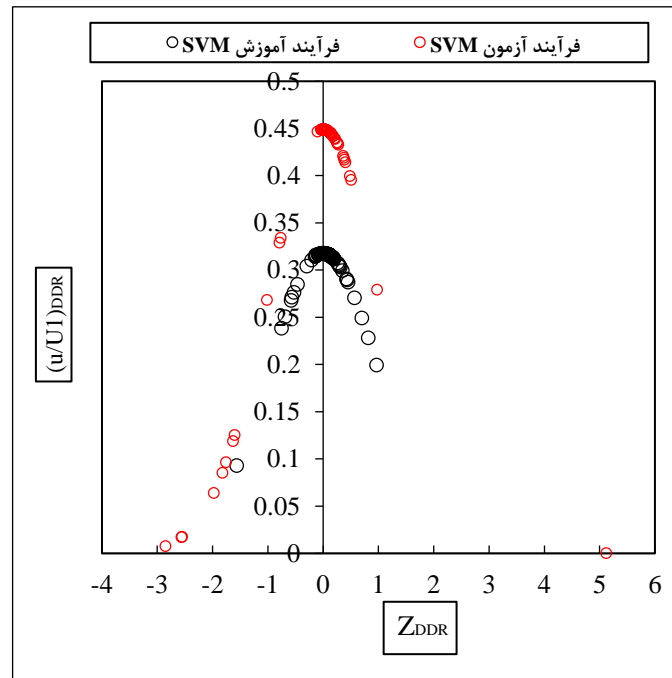
علاوه بر شاخص‌های ارزیابی عملکرد R^2 و RMSE، مقایسه فازهای آموزش و آزمون الگوریتم Nu -SVM به کمک شاخص $(\frac{u}{u_1})_{DDR}$ در جدول (۵) ارائه شده است.

جدول (۵): شاخص‌های ارزیابی عملکرد الگوریتم Nu -SVM در

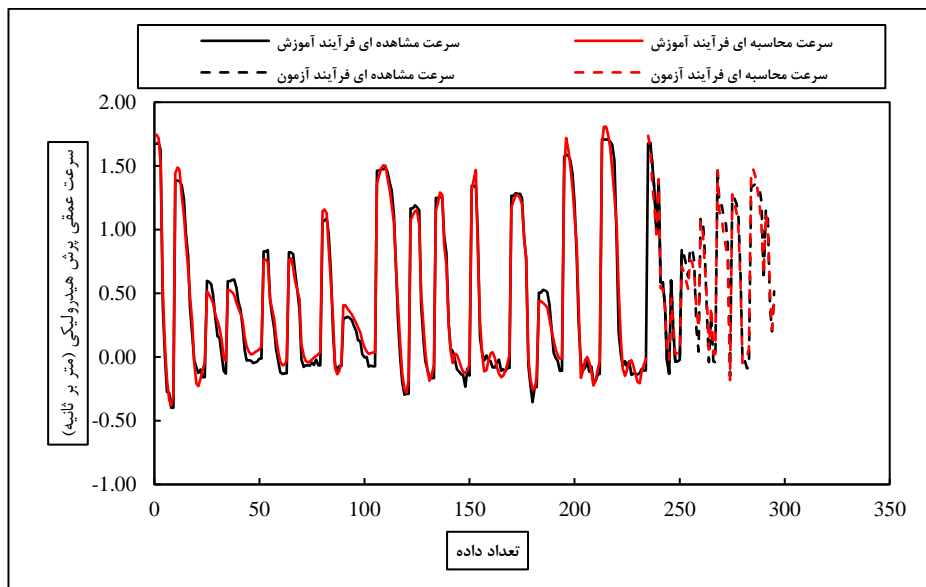
پیش‌بینی سرعت عمقی

الگوریتم	فاز	R^2	RMSE	$(\frac{u}{u_1})_{DDR(max)}$
Nu-SVM	آموزش	۰/۹۸۲۵	۰/۰۸۳۷۶	۰/۳۱۷۸
	آزمون	۰/۹۷۷۰	۰/۰۹۵۸۸	۰/۴۴۸۹

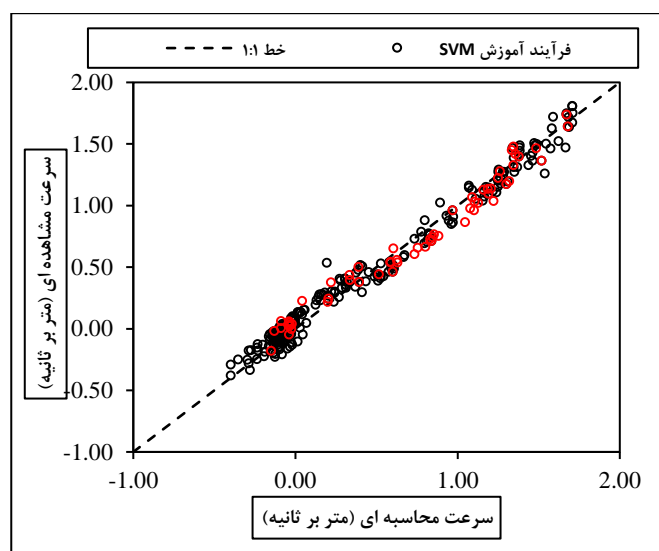
همانطور که مشاهده می‌شود مقادیر RMSE و R^2 برای فرآیندهای آموزش و آزمون تقریباً به هم نزدیک هستند. اما مقایسه‌ی کارکرد الگوریتم براساس شاخص $(\frac{u}{u_1})_{DDR}$ (شکل (۷)) نشان از برتری الگوریتم در فرآیند آزمون دارد



شکل (۷): توزیع Z_{DDR} برای الگوریتم SVM طی فرآیندهای آموزش و آزمون



شکل (۸): سرعت عمقی مشاهده ای و محاسبه ای با الگوریتم SVM در فرآیندهای آموزش و آزمون



شکل (۹): توزیع سرعت‌های محاسبه‌ای و مشاهده‌ای در الگوریتم Nu-SVM

جدول (۶): فراسنجه‌های نهایی مورد استفاده در پیش‌بینی

توزیع عمقی سرعت در الگوریتم GEP

مقدار	پارامتر
۷	اندازه‌ی سر
۳۱	تعداد کروموزوم
۳	تعداد ژن
۰/۰۳۸	نرخ جهش
۰/۱	نرخ وارون‌سازی
۰/۳	نرخ ترکیب تک
۰/۳	نرخ ترکیب دو
۰/۱	نرخ ترکیب ژن
۰/۱	نرخ ترانهش درج متوالی
۰/۱	نرخ ترانهش ریشه درج
۰/۱	نرخ ترانهش ژن
RMSE	معیار خطای تابع برازش
+	تابع پیوند

الگوریتم GEP

اجرای الگوریتم GEP روی داده‌های مشاهده‌ای با ترکیبی از عملگرهای ریاضی و ویژگی‌های مختلفی از کروموزوم‌ها همراه است. در جدول (۶) ویژگی‌های کروموزوم‌ها برای الگوریتم GEP ارائه شده‌اند. هر پنج پارامتر ذکر شده در معادله‌ی (۱۲) برای پیش‌بینی سرعت عمقی در این مدل مورد استفاده قرار گرفتند. برای مدل‌سازی با استفاده از الگوریتم GEP، ۷۵ درصد داده‌ها برای فرآیند آموزش و ۲۵ درصد داده‌ها برای آزمون الگوریتم به کار گرفته شدند و منجر به جواب بهینه گردیدند.

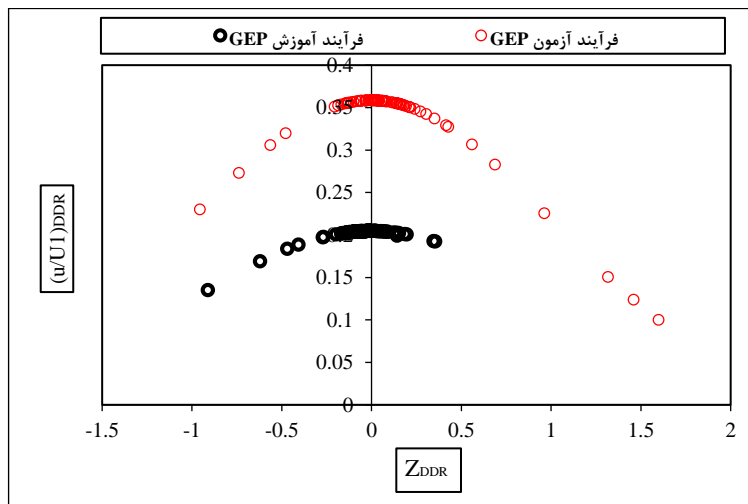
در جدول (۷) خلاصه‌ی مقادیر شاخص‌های ارزیابی عملکرد الگوریتم GEP در فرآیندهای آموزش و آزمون ارائه شده است. همانطور که شاخص‌های عملکرد R^2 و RMSE نشان می‌دهند عملکرد مدل در فرآیند آزمون بهتر از فرآیند آموزش است.

جدول (۷): شاخص‌های ارزیابی عملکرد الگوریتم GEP در پیش‌بینی سرعت عمقی

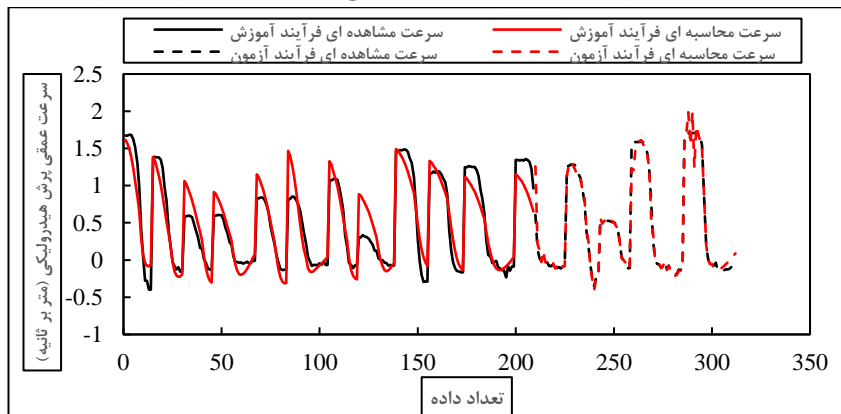
الگوریتم	فاز	R^2	RMSE	$(\frac{u}{U_1})_{DDR}$
GEP	آموزش	۰/۸۶۰۹	۰/۲۱۸۳	۰/۲۰۴۵
	آزمون	۰/۹۷۱۸	۰/۱۱۶۱	۰/۳۵۸۸

شکل (۱۱) مقدار سرعت عمقی پرش هیدرولیکی در طی فرآیندهای آموزش و آزمون الگوریتم GEP برای تمام داده‌ها رسم شده است. در شکل (۱۲) پراکنش سرعت‌های محاسبه‌ای و مشاهده‌ای حول خط ۱:۱ در الگوریتم GEP نشان داده شده است. پراکنش داده‌های فرآیند آزمون حول خط ۱:۱ بهتر از فرآیند آموزش است که نشان از آموزش بهینه مدل می‌باشد. در شکل (۱۳) بیان درختی از الگوریتم GEP همراه با عملگرها و توابع ریاضی نشان داده شده است.

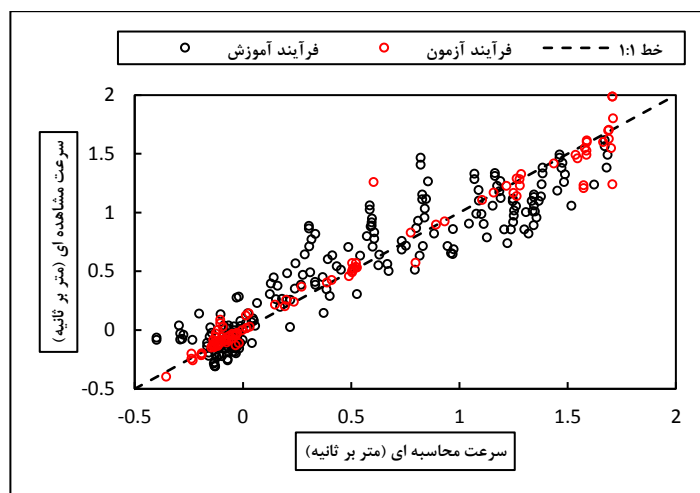
در شکل (۱۰) توزیع خطای به دست آمده از شبیه‌سازی الگوریتم GEP براساس مقدار تابع گوس استاندارد شده ارائه شده است. مقدار بیشینه‌ی $(\frac{u}{U_1})_{DDR}$ در فرآیند آموزش و آزمون به ترتیب ۰/۲۰۴۵ و ۰/۳۵۸۸ است که حاکی از برتری شبیه‌سازی طی فرآیند آزمون است. در



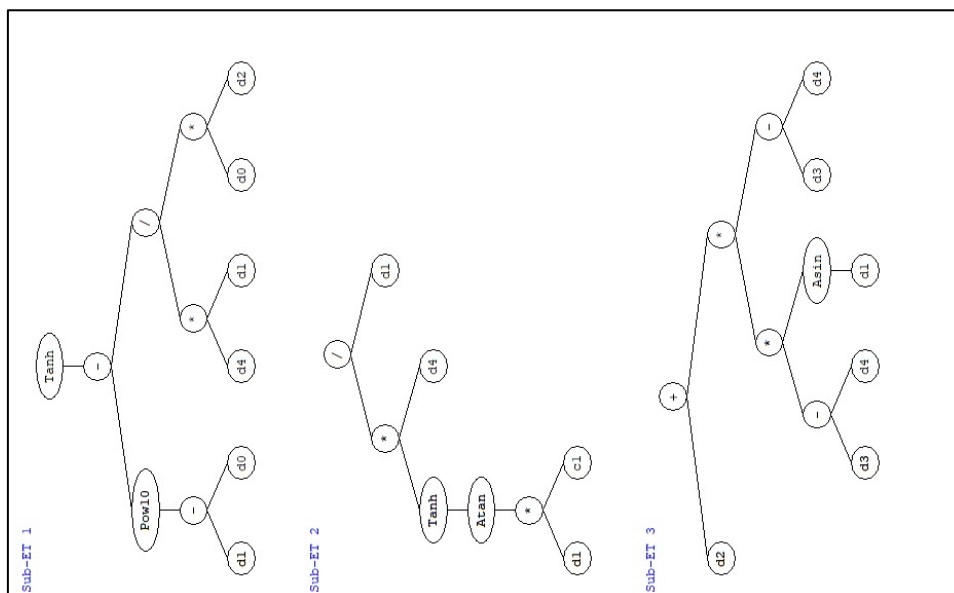
شکل (۱۰): توزیع Z_{DDR} برای الگوریتم GEP طی فرآیندهای آموزش و آزمون



شکل (۱۱): سرعت عمقی مشاهده‌ای و محاسبه‌ای با الگوریتم GEP در فرآیندهای آموزش و آزمون



شکل (۱۲): توزیع سرعت محاسبه‌ای و مشاهده‌ای در الگوریتم GEP



شکل (۱۳): بیان درختی خروجی الگوریتم GEP برای پیش‌بینی توزیع عمقی سرعت در پرش هیدرولیکی مستغرق

معادله‌ی متناظر با بیان درختی شکل (۱۳) به قرار زیر

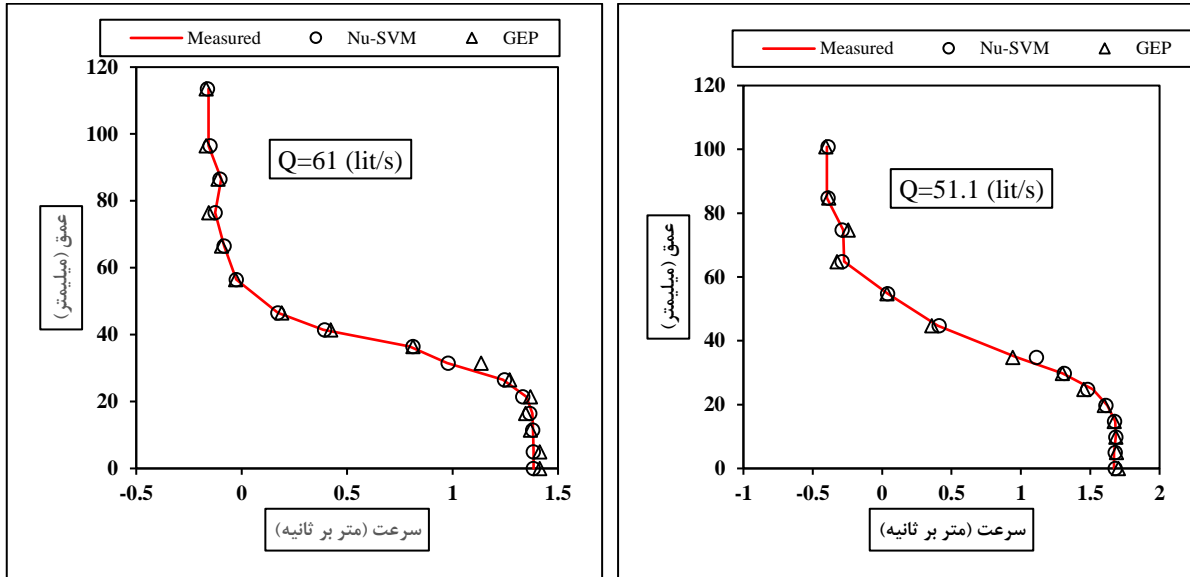
است:

$$\frac{u}{U_1} = \left[\text{Tanh} \left[\left(d_1 - d_2 \right)^{10} - \frac{(d_4 * d_1)}{(d_0 * d_2)} \right] \right] + \left[\tanh(\text{atan } d_1 * c_1) \right] \frac{d_4}{d_1} + \left[d_2 + \left[\left((d_3 - d_4) * \text{Asin } d_1 \right) * (d_3 - d_4) \right] \right] \quad (18)$$

رسم شده‌اند. مقدار زوج مرتب (R^2, RMSE) به ازای دبی ۵۱/۱ لیتر بر ثانیه در الگوریتم‌های SVM و GEP به ترتیب (۰/۹۹۹۸۶۸، ۰/۰۰۷۶۲۴) و (۰/۹۹۶۱۷۲، ۰/۰۴۵۹۱۲) به دست آمده‌اند. مقدار زوج مرتب فوق برای دبی ۶۱ لیتر بر ثانیه برای الگوریتم‌های SVM و GEP به ترتیب (۰/۹۹۹۹۵۸، ۰/۰۰۵۵۹) و (۰/۹۹۸۹۱، ۰/۰۲۹۹۶۳)

که در این معادله d_0, d_1, d_2, d_3, d_4 به ترتیب متغیرهای $Fr_1, Fr_3, \frac{y_1}{y_3}$ و $\frac{w}{b}$ هستند. همچنین مقدار ضریب ثابت c_1 برابر با ۹/۳۹۲۲ می‌باشد. در شکل (۱۴) توزیع سرعت عمقی شبیه‌سازی شده توسط الگوریتم‌های SVM و GEP در مقایسه با مقدارهای مشاهده‌ای رسم شده است. در این شکل، پروفیل عمقی سرعت در پرش هیدرولیکی مستغرق به ازای دو دبی ۵۱/۱ و ۶۰ لیتر بر ثانیه

هستند. مقادیر به دست آمده برای سنج‌های ارزیابی نشان از برتری نسبی الگوریتم SVM نسبت به GEP دارد.



شکل (۱۴): مقایسه‌ی بین سرعت‌های مشاهده‌ای و محاسبه‌ای الگوریتم‌های SVM و GEP به ازای دبی‌های مختلف

نتیجه‌گیری

در این پژوهش از دو الگوریتم Nu-SVM و GEP برای مدل‌سازی توزیع عمقی سرعت در پرش هیدرولیکی مستغرق استفاده شده است. بدین منظور ۳۱۲ داده آزمایشگاهی اندازه‌گیری شده برای انجام پژوهش به کار گرفته شدند. با استفاده از تحلیل ابعادی، پنج پارامتر بی‌بعد موثر در توزیع عمقی سرعت در قالب پارامترهای Fr_3 ، Fr_1 ، $\frac{W}{b}$ و $\frac{Z}{b} \frac{y_1}{y_3}$ استخراج شدند. آزمون Γ ، استفاده از هر پنج پارامتر را به عنوان سنج‌های موثر در مدل‌سازی توزیع عمقی سرعت تایید کرد. بعد از نرمال‌سازی داده‌ها، هر یک از دو الگوریتم Nu-SVM و GEP روی داده‌های نرمال اجرا شدند. در الگوریتم SVM، ۸۰ درصد داده‌ها برای آموزش و ۲۰ درصد داده‌ها برای آزمون مورد استفاده قرار گرفتند. مطلوب‌ترین پاسخ برای الگوریتم SVM به ازای مقادیر ۷ و ۱/۲ به ترتیب برابر با ۰/۴۸۶ و ۰/۴۸۶ با تابع کرنل از نوع RBF به دست آمد. در الگوریتم SVM اگرچه شاخص‌های ارزیابی عملکرد RMSE و R^2 در هر دو فرآیند آموزش و آزمون تقریباً با یکدیگر برابر بودند اما بیشینه مقدار شاخص

در این پژوهش از دو الگوریتم Nu-SVM و GEP برای مدل‌سازی طی فرآیند آموزش با مقدار متناظر $\frac{u}{U_1}$ DDR برای فرآیند آزمون با مقدار ۰/۴۴۸۹ نشان از دقت بهتر مدل‌سازی طی فرآیند آموزش با مقدار متناظر $\frac{u}{U_1}$ DDR داشت. الگوریتم GEP نیز برای همان داده‌های آزمایشگاهی مورد استفاده قرار گرفت اما با این تفاوت که ۷۵ درصد داده‌ها برای آموزش و ۲۵ درصد مابقی برای آزمون الگوریتم استفاده شدند. مقادیر به دست آمده برای سنج‌های R^2 ، RMSE نشان داد عملکرد الگوریتم GEP در فرآیند آزمون بهتر از فرآیند آموزش است که این مطلب ضمن تایید قابلیت GEP در شبیه‌سازی سرعت عمقی جریان، تایید کننده درستی فرآیند صحت‌سنجی مدل نیز می‌باشد. بیشینه مقدار $\frac{u}{U_1}$ DDR طی فرآیند آزمون الگوریتم GEP برابر با ۰/۳۵۸۸ حاصل شد در حالی که مقدار متناظر آن برای فرآیند آموزش ۰/۲۰۴۵ حاصل شد. حال مقایسه‌ی بین مقدار شاخص‌های ارزیابی عملکرد RMSE، R^2 و $\frac{u}{U_1}$ DDR طی فرآیند آزمون الگوریتم‌های SVM و GEP نشان داد الگوریتم SVM از دقت نسبی

بیشتری نسبت به الگوریتم GEP در پیش‌بینی سرعت عمقی جریان در پرش هیدرولیکی مستغرق برخوردار است.

منابع

- پورعبدالله، ن.، حیدرپور، م.، عابدی‌کوپایی، ج. و ج. محمدزاده‌هابیلی. ۱۳۹۹. اثر شیب منفی، زبری بستر و پله انتهایی بر پروفیل سرعت و نوسانات فشار پرش هیدرولیکی. مجله علوم و مهندسی آبیاری، دوره چهل و سه، شماره ۱، ص ۱۷۳-۱۸۸.
- پورعبدالله، ن.، هنر، ت. و ر. فتاحی. ۱۳۹۴. بررسی پروفیل سطح آب و سرعت در پرش هیدرولیکی بر روی بستر زبر با شیب معکوس. نشریه دانش آب و خاک، جلد بیست و پنج، شماره ۱، ص ۱۴۳-۱۵۲.
- تبریزی، ر. و س.م. ثاقبیان. ۱۳۹۸. بررسی مقایسه‌ای پارامترهای هیدرولیکی تأثیرگذار بر مشخصات پرش هیدرولیکی در کانال‌های با مقاطع مرکب با استفاده از روش مبتنی بر کرنل SVM. نشریه آبیاری و زهکشی ایران، شماره چهار، جلد ۱۳، ص ۹۸۷-۹۹۷.
- روشنگر، ک. و ر. ولی‌زاده. ۱۳۹۶. تخمین مشخصات هیدرولیکی کانال‌های واگرای مستهلک کننده انرژی با استفاده از روش ماشین بردار پشتیبان. نشریه علوم آب و خاک (علوم و فنون کشاورزی و منابع طبیعی)، سال بیست و یک، شماره ۲، ص ۲۰۵-۲۱۹.
- شرفی‌گرم‌دره، ا.، وفاه‌خواه، م. و س.س. اسلامیان. ۱۳۹۸. ارزیابی کارایی سامانه‌های ماشین بردار پشتیبان و شبکه عصبی مصنوعی در تحلیل منطقه‌ای سیلاب (مطالعه موردی: حوضه آبخیز دریاچه نمک). نشریه علوم آب و خاک، سال بیست و سه، شماره ۱، ص ۳۵۱-۳۶۶.
- عباسپور، ا. ۱۳۹۳. پیش‌بینی مشخصات پرش هیدرولیکی بر روی بستر زبر با استفاده از شبکه عصبی مصنوعی و برنامه-ریزی ژنتیک. نشریه دانش آب و خاک، جلد بیست و چهار، شماره ۲، ص ۱-۱۰.
- عباسپور، ا.، ع. حسین‌زاده دلیر، د. فرسادی‌زاده و ع.ا. صدرالدینی. ۱۳۸۸. تأثیر بستر موج‌دار سینوسی بر خصوصیات پرش هیدرولیکی. مجله دانش آب و خاک، جلد نوزده، شماره ۱، ص ۱۳-۲۶.
- علیزاده، ا.، ف. یوسفوند و ا. رجبی. ۱۳۹۸. مدل‌سازی طول پرش هیدرولیکی بر روی بسترهای زبر شیب‌دار با استفاده از سیستم استنباط فازی عصبی تطبیقی-الگوریتم ژنتیک. نشریه دانش آب و خاک، جلد بیست و نه، شماره ۴، ص ۱۸۹-۲۰۱.
- غزالی، م.، صمدی‌بروجنی، ح.، قربانی، ب. و ا. رحمتی. ۱۳۹۱. مطالعه آزمایشگاهی پروفیل سرعت در طول پرش هیدرولیکی بر روی بستر با زبری‌های مثلثی. مجله مهندسی آبیاری و آب، سال دو، شماره ۸، ص ۱۱۷-۱۲۸.
- فرزین، س. و م. ولیخان. ۱۳۹۹. بهبود عملکرد ماشین‌های یادگیری در تخمین و پیش‌بینی ضریب آبگذری سرریز. مجله آبیاری و آب ایران، دوره یازده، شماره ۱، ص ۱-۱۳.
- فولادی‌پناه، م.، م. ماجدی‌اصل و آ. حق‌گویی. ۱۳۹۹. کاربرد الگوریتم‌های هوشمند برای مدل‌سازی رابطه دبی-اشل در شرایط استغراق سرریزهای کنگره‌ای و خطی. نشریه هیدرولیک، دوره پانزده، شماره ۲، ص ۱۵۱-۱۶۴.
- ماجدی‌اصل، م. و م. فولادی‌پناه. ۱۳۹۷. کاربرد روش‌های تکاملی در تعیین ضریب دبی سرریزهای کنگره‌ای مثلثی. نشریه علوم آب و خاک، دوره بیست و دو، شماره ۴، ص ۲۷۹-۲۹۰.
- نیک‌پور، م.ر. و س. محمودی‌بابلان. ۱۳۹۷. مقایسه‌ی کارایی مدل‌های هوشمند در روندیابی جریان روزانه‌ی رودخانه (مطالعه‌ی موردی: رودخانه بالخلوچای، استان اردبیل). مجله آبیاری و آب ایران، دوره هشت، شماره ۴، ص ۶۴-۷۸.
- هنر، ت. و س. پورحمزه. ۱۳۹۱. کاربرد شبکه عصبی مصنوعی در برآورد مشخصات پرش هیدرولیکی در حوضچه آرامش با دیواره همگرا و شیب کف منفی. نشریه هیدرولیک، سال هفت، شماره ۱، ص ۸۳-۹۱.

هوشیاری پور، ف.، دهقان، م. و س.ح. مهاجری. ۱۳۹۸. بررسی تاثیر زاویه واگرایی حوضچه آرامش و موقعیت آب پایه بر خصوصیات پرش هیدرولیکی با مدل سازی عددی. مجله مهندسی عمران و محیط زیست امیرکبیر، سال پنجاه و یک، شماره ۱، ص ۸۵-۹۸.

Bijankhan, M., S. Kouchakzadeh and G. Belaud. 2017. Application of the submerged experimental velocity profiles for the sluice gate's stage-discharge relationship. *Flow Measurement and Instrumentation*, 54: 97-108.

Gu, S., Bo F., M. Luo, E. Kazemi, Y. Zhang and J. Wei. 2019. SPH Simulation of Hydraulic Jump on Corrugated Riverbeds. *Applied science*, 9(436): 1-18.

Houichi, L., N. Dechemi, S. Heddami and B. Achour. 2013. An evaluation of ANN methods for estimating the lengths of hydraulic jumps in U-shaped channel. *Journal of Hydroinformatics*, 15(1): 147-154.

Jesudhas, V., V. Roussinova, R. Balachandar and R. Barron. 2017. Submerged Hydraulic Jump Study Using DES. *Journal of hydraulic engineering*, 143(3): 1-11.

Naseri, M. and F. Othman. 2012. Determination of the length of hydraulic jumps using artificial neural networks. *Advances in Engineering Software*, 48: 27-31.

Noori, R., A. Khakpour, B. Omidvar, and A. Farokhnia. (2010). Comparison of ANN and principal component analysis-multivariate linear regression models for predicting the river flow based on developed discrepancy ratio statistic. *Expert Systems with Applications*, 37: 5856-5862.

Noori, R., A.R. Karbassi, A. Moghaddamnia, D. Han, M.H. Zokaei-Ashtiani, A. Farokhnia and M. Ghafari. 2011. Assessment of input variables determination on the SVM model performance using PCA, Gamma test and forward selection techniques for monthly stream flow prediction. *Journal of Hydrology* 401(3-4): 177-189.

Rajaratnam, N. and K. Subramanya. 1967. Flow immediately below a submerged sluice gate, *Journal of Hydraulic Division*, 93 (HY4): 57-77.

Saghebian, S.M. and D. Dragomir-Stanciu. 2019. Hydraulic jump energy dissipation evaluation in rough bed channels via kernel-based approach. 2nd Cilicia International Symposium on Engineering and Technology, 10-12 October, Mersin, Turkey.

Vapnik. V. (1995). *The Nature of Statistical Learning Theory*. Springer-Verlag. New York. 187P.



Performance Assessment of Various Data-Mining Methods to Determine Depth Velocity Profile at Submerged Hydraulic Jump

Mahdi Majedi-Asl^۱, Mehdi Fuladipanah^۲

Abstract

Vertical velocity distribution at hydraulic jump is one of the challenging and significant issues among researchers because of the complexity of measurement and calculations. In this research, the application of the SVM and GEP intelligent models has been considered to determine the vertical velocity profile at the submerged hydraulic jump downstream of a sluice gate. Laboratory measured data of number 312 has been used in the simulation. Using dimensional analysis, dimensionless input parameters were introduced to models including upstream Froud number (Fr_1), Tail water Froud number (Fr_3), the ratio of upstream flow depth to the tailwater depth ($\frac{y_1}{y_3}$), the ratio of the gate opening to the channel width ($\frac{w}{b}$), and the ratio of vertical distance from the channel bed to the channel width ($\frac{z}{b}$). Using the gamma test, all five parameters were determined as the optimum combination to simulate velocity profile. Of two Nu-SVM and C-SVM classification models, the first one was opted as optimum model of the SVM algorithm with RBF Kernel function with the setting parameters γ and Nu of values 1.2 and 0.486, respectively. The performance of the Nu-SVM and the GEP intelligent models were assessed using statistical criteria. The results showed that the values of (RMSE, R^2 , $(\frac{u}{U_1})_{DDR(max)}$) indices for the test phase of the Nu-SVM and the GEP algorithms are (0.09588, 0.9770, 0.4489) and (0.1161, 0.9718, 0.3588) respectively; illustrating the superiority of the Nu-SVM algorithm. Also, according to the gana test, the arrangement of the effective dimensionless parameters on the velocity profile is $\frac{z}{b}$, Fr_1 , $\frac{w}{b}$, Fr_3 and $\frac{y_1}{y_3}$.

Keywords: Flow Velocity, Gamma test, Hydraulic jump, Gene Expression Programming, Support Vector Machine.

^۱ Associate Professor, Department of Civil Engineering, University of Maragheh, Maragheh, Iran (Corresponding Author). mehdi.majedi@gmail.com

^۲ Assistant Professor, Department of Civil Engineering, Ramhormoz Branch, Islamic Azad University, Ramhormoz, Iran. fuladipanah@gmail.com