



# Hybrid Model of Least Square-Support Vector Machine and Particle Swarm Optimization for Prediction of Pavement International Roughness Index

Morteza Araghi<sup>1\*</sup>, Mohsen Khatibinia<sup>2</sup>, Sadegh Moodi<sup>3</sup>

<sup>1</sup>Assistant Professor, Department of Civil Engineering, Faculty of Engineering, University of Birjand, Birjand, Iran.

<sup>2</sup>Associate Professor, Department of Civil Engineering, Faculty of Engineering, University of Birjand, Birjand, Iran.

<sup>3</sup>Assistant Professor, Department of Civil Engineering, Faculty of Mining, Civil and Chemical Engineering, Birjand University of Technology, Birjand, Iran.

## ARTICLE INFO

### Article Type:

Original Research

**Received:** 12.01.2025

**Revised:** 05.03.2025

**Accepted:** 22.04.2025

### Keyword:

Road pavement

International Roughness Index

LTPP database

Least Squares-Support Vector Machine

Particle Swarm Optimization

K-fold Cross Validation

### \*Corresponding Author:

Morteza Araghi

**Email:** [m.araghi@birjand.ac.ir](mailto:m.araghi@birjand.ac.ir)

## ABSTRACT

For the successful management of highway systems, the prediction of pavement performance is vital because it enables distinction and prioritizes maintenance and renovation activities. A precise assessment is necessary for informed and appropriate decisions regarding the repair, refining, and reconstruction of roads and highways. This assessment considers the standards of service and performance of pavements over the years of operation. The International Roughness Index (IRI) has been widely used as an efficient metric for measuring and assessing surface roughness. In this study, a hybrid model was developed for the prediction of IRI. The model was proposed by combining the Least Squares Support Vector Machine (LS-SVM) and Particle Swarm Optimization (PSO). PSO determines the optimal tuning parameters for the LS-SVM in the proposed model. Moreover, the k-fold cross-validation was used during the training phase to avoid overfitting phenomenon. Two datasets from the Long-Term Pavement Performance program were used to assess the performance of the combined model in predicting IRI, and two datasets from the Long-Term Pavement Performance (LTPP) program were used. For the two LTPP datasets, the proposed hybrid model achieved the higher accuracy in term of R<sup>2</sup> values of 0.994 and 0.997, respectively. In addition, a comparison of the proposed model with different machine-learning methods confirmed that the proposed model provided more accurate prediction.



---

## EXTENDED ABSTRACT

---

### Introduction

Roads play a crucial role in transportation and society. To preserve the integrity of the road network, consistent maintenance efforts are necessary. This requires making decisions that are both technically sound and economically viable [1]. A considerable portion of government infrastructure budgets is typically directed towards the expenses associated with these activities. Consequently, restricting investment has detrimental effects on the road network [2-4]. Therefore, it is crucial to optimize funding to maintain the quality of the road infrastructure.

The transportation sector is undergoing significant changes due to advances in artificial intelligence, big data, autonomous vehicles, and carbon reduction. However, these advances pose technical challenges to reevaluate traditional paradigms [5; 6]. Currently, artificial intelligence has influenced numerous sectors, with machine learning techniques being particularly noteworthy.

To guarantee safe and properly maintained roads for users, evaluating the performance of road pavements is essential. There are various methods available for assessing pavement performance, with some of the more commonly used is the International Roughness Index (IRI), Pavement Condition Index (PCI), and Present Serviceability Index (PSI). The IRI quantitatively reflects the longitudinal profile of pavement and originated from the World Bank's 1982 experiment in Brazil [7].

This study proposes a hybrid model that integrates the Least Squares-Support Vector Machine (LS-SVM) with Particle Swarm Optimization (PSO) for predicting IRI. The tuning parameters for the LS-SVM model are identified through the PSO algorithm. Furthermore, to avert overfitting during the training phase, the k-fold cross-validation method is utilized. To facilitate IRI prediction, two databases are employed for training and testing the model. Ultimately, the effectiveness and capabilities of the proposed hybrid model are illustrated by comparing its results against those from other methodologies featured in earlier studies.

### Methodology

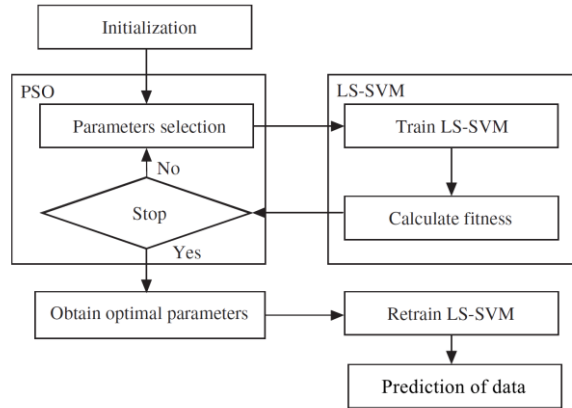
In this study, the tuning parameters and kernel function of the Gaussian Radial Basis Function (GRBF) are determined using the Particle Swarm Optimization (PSO) algorithm. For this purpose, an appropriate objective function must be selected for a set of training data, while ensuring that the risk of overfitting during the optimization process is minimized. This issue can be addressed by employing the cross-validation technique [8]. In this context, k-fold cross-validation provides the best balance between computational cost and reliable parameter estimation [9]. In this method, for each particle in the PSO algorithm, an LS-SVM model is created using the parameters associated with that particle. The training data is divided into k subsets, where k-1 subsets are used as the training set for the LS-SVM model, and the performance of the model is evaluated using the Mean Absolute Percentage Error (MAPE) criterion on the kth subset. This process is repeated k times, ensuring that

each subset is used exactly once for testing. Consequently, minimizing the average MAPE across all k trials is defined as the objective function for each particle, as follows:

$$\text{Minimize: } f = \sum_{i=1}^k \text{MAPE}_i \tag{1}$$

$$\text{MAPE}_i = \frac{\sum_{i=1}^l |a_i - p_i| / a_i}{l}, \quad i = 1, 2, \dots, k \tag{2}$$

where  $a_i$  and  $p_i$  are the actual value and the predicted value, respectively. Figure 1 represents the number of test data in the kth subset. Figure 1 shows the flowchart of the combined LS-SVM and PSO models.



**Figure 1** Flowchart of the combined LS-SVM and PSO model

The performance of the combined LS-SVM and PSO model in predicting IRI was evaluated in terms of three statistical measures including Mean Absolute Percentage Error; Root Mean Square Error (*RMSE*), and coefficient of determination ( $R^2$ ). These *RMSE* and  $R^2$  measures are expressed as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (a_i - p_i)^2}{n}} \tag{2}$$

$$R^2 = \frac{\left( \sum_{i=1}^n (a_i - a_{ave})(p_i - p_{ave}) \right)^2}{\sum_{i=1}^n (a_i - a_{ave})^2 \sum_{i=1}^n (p_i - p_{ave})^2} \tag{3}$$

where  $a_{ave}$  and  $p_{ave}$  are the average of the actual and predicted values of the sample in the databases.

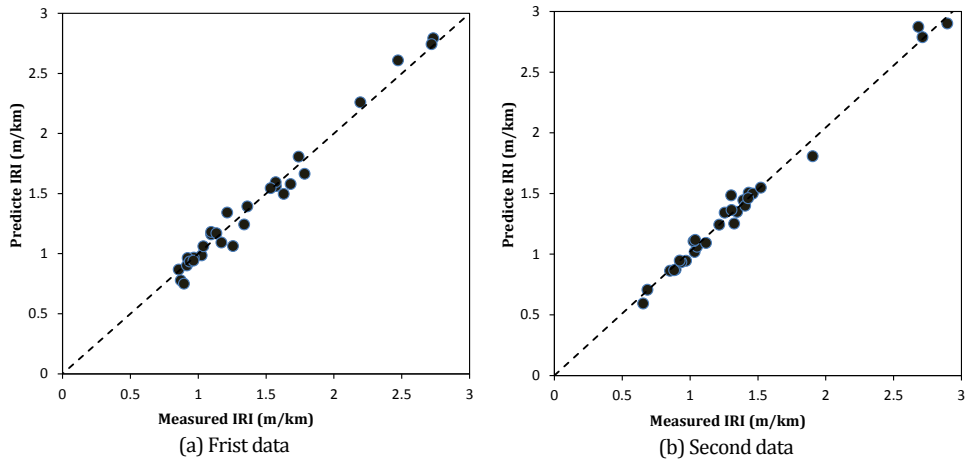
### Results and discussion

Table 1 shows the prediction performance of the combined LS-SVM and PSO model for the training and test sets of the two databases. The comparison of the data shows that the combined LS-SVM and PSO model has been able to predict the IRI in the two databases with very high accuracy.

**Table 1. Performance of the combined LS-SVM and PSO model in IRI prediction**

Data	Train			Test		
	MAPE	RMSE	$R^2$	MAPE	RMSE	$R^2$
First	1.187	0.012	0.999	5.126	0.055	0.997
Second	4.334	0.057	0.996	3.917	0.074	0.994

The scatter plots between the measured (actual) and predicted IRI values in the testing process for both databases are shown in Figure 3. The results in Figure 2 show that the proposed hybrid model was able to predict the IRI value well.

**Figure 2. Scatter plots between measured and predicted IRI values**

## Conclusion

In this study, a combined LS-SVM and PSO model was proposed for IRI prediction. The PSO algorithm was used to optimally determine the tuning parameters of the LS-SVM model. Additionally, the k-fold cross-validation technique was employed to prevent overfitting during the training process. Two laboratory databases from the Long-Term Pavement Performance (LTPP) program were used to evaluate the efficiency of the proposed combined model. The accuracy evaluation of the combined LS-SVM and PSO model demonstrated high performance, with coefficients of determination ( $R^2$ ) of 0.994 and 0.997 for the two LTPP databases. These results indicate the model's high accuracy in IRI prediction. Furthermore, the performance of the combined model was compared with other machine learning methods in terms of the correlation coefficient (R). The results revealed that the combined LS-SVM and PSO models achieved superior accuracy. Therefore, the proposed model can be effectively utilized for pavement performance prediction.



## مدل ترکیبی ماشین بردار پشتیبان حداقل مربعات و بهینه‌سازی ازدحام ذرات برای پیش‌بینی شاخص بین‌المللی ناهمواری روسازی

مرنضی عراقی<sup>۱\*</sup>، محسن خطیبی‌نیا<sup>۲</sup>، صادق مودی<sup>۳</sup>

۱- استادیار، گروه مهندسی عمران، دانشکده مهندسی، دانشگاه بیرجند، بیرجند، ایران.

۲- دانشیار، گروه مهندسی عمران، دانشکده مهندسی، دانشگاه بیرجند، بیرجند، ایران.

۳- استادیار، گروه مهندسی عمران، دانشکده مهندسی معدن، عمران، شیمی، دانشگاه صنعتی بیرجند، بیرجند، ایران.

### چکیده

### اطلاعات مقاله

نوع مقاله: مقاله پژوهشی

دریافت مقاله: ۱۴۰۳/۱۰/۲۳

بازنگری مقاله: ۱۴۰۳/۱۲/۱۵

پذیرش مقاله: ۱۴۰۴/۰۲/۰۲

### کلید واژگان:

روسازی جاده‌ها

شاخص بین‌المللی ناهمواری

پایگاه داده LTTP

ماشین بردار پشتیبان حداقل مربعات

بهینه‌سازی ازدحام ذرات

اعتبارسنجی متقاطع k-fold

پیش‌بینی عملکرد روسازی برای مدیریت مؤثر زیرساخت‌های جاده‌ای بسیار مهم است. زیرا به شناسایی و اولویت‌بندی فعالیت‌های نگهداری و نوسازی کمک می‌کند. اتخاذ تصمیمات آگاهانه و مناسب در خصوص تعمیر، بهسازی و بازسازی روسازی جاده‌ها و بزرگراه‌ها مستلزم برآورد دقیق کیفیت خدمت‌دهی و عملکرد روسازی در طی سالیان بهره‌برداری است. شاخص بین‌المللی ناهمواری (IRI) به عنوان یک شاخص کارآمد برای ارزیابی و تحلیل ناهمواری‌های سطح روسازی استفاده می‌شود. در این مقاله، مدل ترکیبی ماشین بردار پشتیبان حداقل مربعات (LS-SVM) و بهینه‌سازی ازدحام ذرات (PSO) توسعه داده شده است. در این مدل پیشنهادی، مقادیر بهینه پارامترهای تنظیم‌کننده LS-SVM توسط PSO تعیین می‌شود. همچنین تکنیک اعتبارسنجی متقاطع k-fold برای جلوگیری از بیش‌برازش LS-SVM در مرحله آموزش استفاده شده است. دو پایگاه داده آزمایشگاهی برنامه عملکرد بلندمدت روسازی (LTTP) برای سنجش کارایی مدل ترکیبی پیشنهادی جهت پیش‌بینی IRI در نظر گرفته شده است. ارزیابی دقت مدل ترکیبی LS-SVM و PSO براساس معیار ضریب تعیین ( $R^2$ ) برای دو پایگاه داده برابر ۰/۹۹۴ و ۰/۹۹۷ می‌باشد. در نهایت، مقایسه عملکرد مدل ترکیبی با دیگر روش‌های یادگیری ماشین نشان می‌دهد که مدل ترکیبی LS-SVM و PSO از دقت بالایی برخوردار می‌باشد.

\*نویسنده مسئول: مرنضی عراقی

پست الکترونیکی:

[m.araghi@birjand.ac.ir](mailto:m.araghi@birjand.ac.ir)



## مقدمه

جاده‌ها نقش بسیار مهمی در حمل و نقل و جامعه ایفا می‌نمایند. برای حفظ کیفیت شبکه جاده‌ای، لازم است فعالیت‌های نگهداری منظم صورت گیرد که این مهم نیاز به تصمیماتی دارد که هم از نظر فنی و هم از نظر اقتصادی امکان‌پذیر باشد [۱]. هزینه این فعالیت‌ها اغلب بخش قابل توجهی از بودجه‌های زیربنایی دولت را تشکیل می‌دهد. بنابراین سرمایه‌گذاری محدود، بر شبکه جاده‌ای تأثیر منفی می‌گذارد [۲-۴]. به این ترتیب، بهینه‌سازی بودجه برای اطمینان از کیفیت شبکه جاده ضروری است.

بخش حمل و نقل به دلیل پیشرفت در هوش مصنوعی، ابرداده‌ها، وسایل نقلیه خودران و کاهش کربن، تغییرات زیادی را تجربه می‌کند. با این وجود، این پیشرفت‌ها چالش‌های فنی را برای ارزیابی مجدد پارادایم‌های سنتی ایجاد می‌کنند [۵؛ ۶]. امروزه هوش مصنوعی به زمینه‌های مختلف نفوذ کرده است که روش‌های یادگیری ماشین یکی از برجسته‌ترین آن‌هاست [۱۰-۱۴]. یادگیری ماشین با پیشرفت‌های الگوریتم، در دسترس بودن داده‌ها و کاهش هزینه‌های محاسباتی توسعه بیشتری یافته است [۱۵]. اساساً یادگیری ماشین شامل توانایی رایانه‌ها برای یادگیری نحوه انجام وظایفی مانند پیش‌بینی، طبقه‌بندی، خوشه‌بندی و تشخیص الگو می‌باشد [۱۵]؛ بنابراین، این یادگیری بدون برنامه‌نویسی صریح رخ می‌دهد [۱۶]. یادگیری ماشین در سال‌های اخیر به دلیل توانایی آن در مدل‌سازی روابط پیچیده بین ورودی‌ها و خروجی‌ها برای پیش‌بینی عملکرد روسازی محبوب شده است [۷؛ ۱۶]. الگوریتم‌های یادگیری ماشین توانایی پیش‌بینی جنبه‌های مختلف عملکرد روسازی از جمله ناهمواری، ترک‌خوردگی و شیارشدگی را دارند [۷]. با آموزش مجموعه داده‌های بزرگ روسازی، این الگوریتم‌ها قادر به شناسایی الگوها و روابطی هستند که ممکن است به راحتی برای چشم انسان قابل تشخیص نباشند. این موضوع منجر به بهبود دقت در پیش‌بینی عملکرد روسازی و مدیریت بهتر زیرساخت‌های جاده‌ای می‌شود [۱۶؛ ۱۷].

مدل‌های عملکرد روسازی (PPM) ابزارهای ضروری برای پیش‌بینی عملکرد روسازی‌های جاده‌ای در طول زمان هستند [۱۷؛ ۱۸]. مهندسان روسازی این مدل‌ها را برای مقابله با چالش پیش‌بینی رفتار سازه‌های پیچیده ساخته‌شده از مواد مختلف که به شرایط ترافیکی و محیطی متفاوت پاسخ می‌دهند، توسعه داده‌اند [۳]. دقت، دامنه و داده‌های مورد نیاز PPMها ممکن است متفاوت باشد و داده‌های با کیفیت بالا برای به حداکثر رساندن اثربخشی آن‌ها مهم است. PPMها به سه دسته اصلی مکانیزمی، تجربی و مکانیزمی-تجربی تقسیم می‌شوند [۱۷]. مدل‌های مکانیزمی از نظر ریاضی به مدل‌سازی فیزیک روسازی‌ها می‌پردازند. بنابراین این مدل واکنش روسازی به بار ترافیک را محاسبه می‌کند. از سوی دیگر، مدل‌های تجربی از تجزیه و تحلیل رگرسیون برای شناسایی عواملی مانند ترافیک، آب‌وهوا، سن روسازی و دیگر عوامل موثر بر روسازی استفاده می‌کنند. این مدل‌ها معمولاً از داده‌های مشاهده‌شده برای ایجاد ارتباط بین ورودی‌ها و خروجی‌ها استفاده می‌کنند. بنابراین، PPMهای مبتنی بر یادگیری ماشین، تجربی هستند. مدل‌های مکانیزمی-تجربی پاسخ‌های تنش و کرنش روسازی را از طریق تجزیه و تحلیل مکانیزمی تعیین می‌کنند و سپس آن‌ها را با عملکرد روسازی یا تخریب از طریق تجزیه و تحلیل رگرسیون مرتبط می‌کنند.

برای اطمینان از اینکه جاده‌ها برای کاربران ایمن بوده و به‌خوبی نگهداری می‌شوند، ارزیابی عملکرد روسازی جاده بسیار مهم است. روش‌های مختلفی برای ارزیابی عملکرد روسازی‌های جاده‌ای وجود دارد که برخی از رایج‌ترین آن‌ها شامل شاخص ناهمواری بین‌المللی (IRI) [۲]، شاخص وضعیت روسازی (PCI) [۳] و شاخص خدمات‌پذیری کنونی (PSI) [۴] می‌باشد. IRI از نظر ریاضی پروفیل طولی روسازی را نشان می‌دهد که از آزمایش ۱۹۸۲ برزیل بانک جهانی ریشه

<sup>1</sup> Pavement Performance Models

<sup>2</sup> International Roughness Index

<sup>3</sup> Pavement Condition Index

<sup>4</sup> Present Serviceability Index

می‌گیرد [۱۹]. این شاخص به عنوان یک معیار رایج و پذیرفته‌شده، ناهمواری روسازی را با محاسبه پروفیل طولی متوسط اندازه‌گیری می‌کند که منعکس‌کننده تغییرات سطحی است که باعث ارتعاشات خودرو می‌شود [۱۹]. این اندازه‌گیری بر اساس پاسخ فرضی یک‌چهارم ماشین‌های در حال حرکت با سرعت ۸۰ کیلومتر در ساعت انجام شده است [۲۰]. شاخص PCI در مقیاس عددی بین صفر و ۱۰۰ تغییر می‌کند که مقادیر بالاتر نشان‌دهنده عملکرد بهتر است [۲۰]. PSI روشی برای ارزیابی وضعیت فعلی یک جاده بر اساس مشاهدات بصری است و از صفر (غیر قابل عبور) تا ۵ (عالی) تغییر می‌کند. PSI واریانس شیب را در نظر می‌گیرد و می‌تواند با معیارهای عملکرد ناهمواری مانند IRI مرتبط باشد [۲۱].

در این مطالعه، مدل ترکیبی ماشین بردار پشتیبان حداقل مربعات (LS-SVM)<sup>۱</sup> و بهینه‌سازی ازدحام ذرات (PSO) برای پیش‌بینی IRI پیشنهاد شده است. برای این منظور، پارامترهای تنظیم‌کننده مدل LS-SVM توسط الگوریتم PSO تعیین می‌شود. همچنین برای جلوگیری از بیش‌برازش مدل LS-SVM در پروسه آموزش، تکنیک اعتبارسنجی متقاطع k-fold<sup>۲</sup> استفاده می‌شود. ابتدا برای پیش‌بینی IRI، از دو پایگاه داده آزمایشگاهی برای آموزش و آزمایش مدل استفاده می‌شود. در نهایت کارایی و توانایی مدل ترکیبی پیشنهادی از مقایسه نتایج مدل ترکیبی با دیگر تکنیک‌های ارائه شده در مطالعات گذشته ارائه می‌شود.

### یادگیری ماشین برای پیش‌بینی IRI

در سال‌های گذشته، مطالعات مختلفی اثربخشی الگوریتم‌های یادگیری ماشین را در پیش‌بینی عملکرد روسازی نشان داده‌اند. الگوریتم‌های رایج شامل شبکه عصبی مصنوعی (ANN)، ماشین بردار پشتیبان (SVM)، رگرسیون بردار پشتیبانی (SVR)، تقویت سازگاری (AdaBoost)، جنگل تصادفی (RF)، درختان تصمیم‌گیری تقویت گردان (GBDT) و تقویت دسته‌بندی (CatBoost) هستند. همچنین، چندین برنامه کاربردی با استفاده از ANN برای پیش‌بینی IRI معرفی شده است. به عنوان نمونه، حسین و همکاران [۲۲] از ANN برای پیش‌بینی IRI استفاده نمودند. عبدالعزیز و همکاران [۲۳] یک مدل پیش‌بینی IRI برای روسازی‌های انعطاف‌پذیر با استفاده از تجزیه و تحلیل ANN و MLR توسعه دادند. علاوه بر این زیادا و همکاران [۲۴] اثبات کردند که ANN در پیش‌بینی عملکرد روسازی در مناطق با آب‌وهوای گرم در مقایسه با روش‌های رگرسیون معمولی عملکرد دقیق‌تری دارد. کاربردهای الگوریتم RF برای پیش‌بینی عملکرد روسازی‌ها نتایج امیدوارکننده‌ای را نشان داده است. گونگ و همکارانش [۲۵] استفاده از RF را برای پیش‌بینی IRI پیشنهاد نمودند و نشان دادند که این روش از دقت بیشتری نسبت به رگرسیون خطی برخوردار می‌باشد. ناصری و همکاران [۲۶] با ترکیب الگوریتم RF و الگوریتم بهینه‌سازی نهنگ (WOA) با دقت بیشتر در پیش‌بینی IRI و تحقق بهینه‌سازی نگهداری روسازی کارآمدتر و مقرون به صرفه‌تر از مدل‌های سنتی دست یافتند.

برخی از پژوهشگران نیز مدل‌های مختلف را برای پیش‌بینی عملکرد روسازی مقایسه کردند. شرما و همکارانش [۲۷] پنج مدل از جمله ANN، GBDT، درختان بسیار تصادفی (XRT)، مدل خطی عمومی (GLM) و RF را مقایسه کردند. یافته‌های آن‌ها نشان داد که عملکرد GBDT در مقایسه با دیگر مدل‌ها بهتر می‌باشد. همچنین، زیادا و همکاران [۲۴] عملکرد روسازی آسفالت را در آب‌وهوای گرم مورد مطالعه قرار دادند و هفت عامل اصلی طراحی را مشخص نمودند. آن‌ها تکنیک‌های یادگیری ماشین از جمله SVM، DT، گروه‌های تقویت شده درختان (EBT)، GPR و ANN را با رگرسیون سنتی مقایسه کردند. طبق نتایج آن‌ها، با وجود عوامل محیطی مختلف که بر عملکرد در مناطق گرم در مقابل سرد تأثیر می‌گذاشت، ANN دقیق‌ترین روش نسبت به دیگر روش‌ها بود. علاوه بر این، ناصری و همکارانش [۲۸] چهار الگوریتم SVM، DT، RF و ANN را برای پیش‌بینی IRI بررسی کردند. این مطالعه همچنین

<sup>1</sup> Least Square-Support Vector Machine

<sup>2</sup> Particle Swarm Optimization

<sup>3</sup> cross validation technique

یک تکنیک انتخاب ویژگی ترکیبی با استفاده از الگوریتم بهینه‌سازی ریاضی و رگرسیون نزولی گرادیان تصادفی (AOA-SGDR) را برای ساده‌سازی مجموعه اولیه ۵۸ متغیره معرفی نمودند. در یک مطالعه دیگر، ژانگ و همکاران [۲۹] از روش GBDT برای پیش‌بینی IRI، شیارشدگی، ترک خستگی، ترک عرضی و ترک طولی استفاده کردند، در حالی که عوامل حیاتی برای عملکرد روکش را نیز با پیش‌روکش و ترک عرضی به عنوان عوامل کلیدی تعیین کننده دوام پوشش شناسایی کردند. دامیرچیلو و همکارانش [۳۰] پیش‌بینی IRI برای روسازی‌های آسفالت را از طریق تکنیک‌های یادگیری ماشین بررسی کردند و مشخص نمودند که افزایش گرادیان شدید (XGBoost) بهترین مدل از نظر عملکرد می‌باشد. سونگ و همکاران [۳۱] یک مدل یادگیری گروهی بر اساس ماشین‌های تقویت گرادیان رعد و برق (ThunderGBM) برای پیش‌بینی IRI، روسازی‌های انعطاف‌پذیر را پیشنهاد کردند. یافته‌های این مطالعه نشان داد که مدل پیشنهادی آن‌ها از مدل‌های ANN، MEPDG، و RF بهتر عمل کرده است. اخیراً، تاماگوسکو و فریرا [۱] در یک مطالعه مروری تمامی مدل‌هایی را که در پیش‌بینی IRI مورد استفاده قرار گرفته است را مورد بررسی قرار داده‌اند.

### ماشین بردار پشتیبان حداقل مربعات (LS-SVM)

ماشین بردار پشتیبان نوعی سیستم یادگیری کارآمد بر مبنای تئوری بهینه‌سازی مقید است که توسط واپنیک و کورتس [۳۲] در سال ۱۹۹۵ ارائه شد. در این مدل از اصل استقرای کمینه‌سازی خطای ساختاری استفاده شده و در بیشتر مسائل به یک جواب بهینه کلی منجر می‌شود. در این مدل برای حل معادلات از برنامه‌ریزی دوگانه استفاده می‌شود. بنابراین حل مسائل در ابعاد بزرگ با استفاده از مدل یادشده، هزینه محاسباتی را به شدت افزایش می‌دهد. برای رفع این مشکل، سویکنز و همکاران [۳۳] روش ماشین بردار پشتیبان حداقل مربعات را ارائه نمودند. در این مدل از معادلات خطی ساده برای حل مسائل استفاده می‌شود که به طور مؤثری پیچیدگی الگوریتم را کاهش می‌دهد. در مدل LS-SVM با فرض معلوم بودن  $N$  داده جهت آموزش  $\{x_k, y_k\}$  با داده‌های ورودی  $x_k \in \mathbb{R}^n$  و خروجی  $y_k \in \mathbb{R}^n$  مسئله بهینه‌سازی زیر در فضای وزن‌دار اولیه به صورت زیر قابل ملاحظه است [۳۳]:

$$\min J(w, e) = \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{k=1}^N e_k^2 \quad (1)$$

به طوری که:

$$y_k (w^T \phi(x_k) + b) = 1 - e_k, \quad k = 1, \dots, N \quad (2)$$

در فضای وزن‌دار اولیه، مدل LS-SVM زیر برای تخمین توابع غیرخطی به صورت زیر است:

$$y(x) = w^T \phi(x) + b \quad (3)$$

برای محاسبه  $w$  و  $e$  روش ضرایب لاگرانژین برای حل معادله (۳) به صورت زیر تعریف می‌شود:

$$\zeta(w, b, e; \alpha) = J(w, e) - \sum_{k=1}^N \alpha_k \{w^T \phi(x_k) + b + e_k - y_k\} \quad (4)$$

<sup>1</sup>pre-overlay

<sup>2</sup> Tamagusko and Ferreira

<sup>3</sup>Vapnik and Cortes

<sup>4</sup>Suykens

$\alpha_k \in \square$  ضرایب لاگرانژین هستند که مقادیر پشتیبان نامیده می‌شوند. شرایط بهینه‌سازی با استفاده از شرایط کان-تاکر به صورت زیر تعریف می‌شود:

$$\frac{\partial \zeta}{\partial b} = 0 \rightarrow \mathbf{w} = \sum_{k=1}^N \alpha_k \phi(\mathbf{x}_k) \quad (5)$$

$$\frac{\partial \zeta}{\partial b} = 0 \rightarrow \sum_{k=1}^N \alpha_k = 0 \quad (6)$$

$$\frac{\partial \zeta}{\partial e_k} = 0 \rightarrow \alpha_k = \gamma e_k, \quad k = 1, \dots, N \quad (7)$$

$$\frac{\partial \zeta}{\partial \alpha_k} = 0 \rightarrow \mathbf{w}^T \phi(\mathbf{x}_k) + b + e_k - \mathbf{y}_k = 0, \quad k = 1, \dots, N \quad (8)$$

این شرایط مشابه با شرایط بهینه‌سازی استاندارد ماشین بردار پشتیبان است، به جز برای شرط  $\alpha_k = \gamma e_k$  بنابراین:

$$\omega = \sum_{k=1}^N \alpha_k \phi(\mathbf{x}_k) = \sum_{k=1}^N \gamma e_k \phi(\mathbf{x}_k) \quad (9)$$

با قراردادن رابطه (۹) در رابطه (۳) داریم که:

$$\mathbf{y} = \sum_{k=1}^N \alpha_k \phi(\mathbf{x}_k)^T \phi(\mathbf{x}_k) + b = \sum_{k=1}^N \alpha_k \langle \phi(\mathbf{x}_k)^T, \phi(\mathbf{x}_k) \rangle + b \quad (10)$$

بعد از حذف  $\mathbf{w}$  و  $\mathbf{e}$  مجموعه روابط خطی زیر برای حل  $\alpha$  زیر به دست می‌آید:

$$\begin{bmatrix} \Omega + \frac{1}{\gamma} \mathbf{I} & \mathbf{I}_N \\ \mathbf{I}_N^T & 0 \end{bmatrix} \begin{cases} \alpha \\ b \end{cases} = \begin{cases} \mathbf{y} \\ 0 \end{cases} \quad (11)$$

که در رابطه فوق،  $\Omega$  ماتریس هسته می‌باشد که در آن درایه  $j$  نام به صورت زیر تعریف می‌شود:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j) \quad (12)$$

که در آن  $K(\mathbf{x}_i, \mathbf{x}_j)$  تابع کرنل می‌باشد. همچنین  $\mathbf{I}$  نشان‌دهنده یک ماتریس واحد  $N \times N$  و  $\mathbf{1}_N = [111\dots 1]^T$  است. بنابراین با محاسبه  $b$  از رابطه (۱۱)، مدل LS-SVM برای تخمین تابع به صورت زیر به دست می‌آید:

$$\mathbf{y}(\mathbf{x}) = \sum_{k=1}^N \alpha_k K(\mathbf{x}, \mathbf{x}_k) + b \quad (13)$$

برای تحلیل حساسیت مدل LS-SVM می‌توان با استفاده از روش آزمون و خطا، تابع کرنل (هسته) را انتخاب نمود. براساس مطالعه لین و همکاران [۳۴]، تابع پایه شعاعی گوسی (GRBF) در دو مقوله نسبت به توابع هسته دیگر دارای نقطه قوت بوده و در نتیجه عملکرد مطلوب‌تری در مسائل پیش‌بینی از خود نشان داده است. این تابع به صورت

<sup>1</sup> Gaussian Radial Basis function

غیرخطی نمونه‌ها را به ابعاد فضایی بالاتری نگاشت می‌نماید، بنابراین برخلاف تابع هسته خطی، در حالت‌هایی که روابط بین نمونه‌ها و ویژگی‌ها به‌صورت غیرخطی است، کاربرد خواهد داشت. دومین نقطه قوت، تعداد پارامترهایی است که بر پیچیدگی انتخاب مدل تأثیرگذار می‌باشد. تابع هسته چندجمله‌ای تعداد پارامترهای بیشتری نسبت به تابع شعاعی گوسی دارد. لذا در این تحقیق از تابع پایه شعاعی گوسی استفاده شده است که به‌صورت زیر بیان می‌شود [۳۳]:

$$K(\mathbf{x}, \mathbf{x}_i) = \exp(-\|\mathbf{x} - \mathbf{x}_i\|^2 / 2\sigma^2) \quad (14)$$

و در آن  $\sigma$  سطح نویز گاوسی انحراف استاندارد است.

### الگوریتم بهینه‌سازی ازدحام ذرات (PSO)

الگوریتم PSO در سال ۱۹۹۵ توسط ابرهات و کندلی [۳۵] معرفی شد. الگوریتم PSO در بین الگوریتم‌های مشابه خود بسیار قوی ظاهر شده است و به‌خوبی با توابع غیرخطی، غیرمحدب، پیوسته و ناپیوسته (گسسته) وفق یافته است. از جمله مزایای دیگر الگوریتم، به پیاده‌سازی آسان آن به شکل روابط ریاضی می‌توان اشاره کرد که به راحتی با شرایط قیود و متغیرها در یک حالت خاص منطبق می‌شود. همچنین کارایی این الگوریتم در بسیاری از مسائل بهینه‌سازی مهندسی توسط محققین گذشته به اثبات رسیده است [۸؛ ۹؛ ۳۶].

این الگوریتم از زندگی گروهی پرندگان الهام گرفته شده است که بسیاری از احتیاجات خود از جمله جستجو و یافتن غذا را به صورت دسته‌جمعی و با کمک یکدیگر انجام می‌دهند. در این الگوریتم هر جواب مسئله، یک پرند در فضای جستجو است که ذره نام دارد و به گروه ذرات یا پرندگان گله گفته می‌شود. هر ذره دارای یک مقدار شایستگی است که توسط تابع شایستگی (برازندگی) مسئله به‌دست می‌آید. بر این اساس پرنده‌ای که به غذا نزدیک‌تر است، شایستگی بیشتری دارد. همچنین هر پرند دارای یک بردار سرعت است که جهت حرکت پرند و اندازه سرعت را نشان می‌دهد. در طول فرآیند بهینه‌سازی هر پرند جهت خود را بر اساس بهترین تجربه شخصی خود ( $pbest$ ) و تجربه بهترین ذره در گله ( $gbest$ ) اصلاح می‌کند. از دیدگاه عددی، ذره نام با بردار موقعیت  $\mathbf{x}$  و بردار سرعت  $\mathbf{v}$  در تکرار  $t$ ام با استفاده از روابط (۱۵) و (۱۶) به ترتیب سرعت و موقعیت خود را بهبود می‌بخشد [۳۵]:

$$\mathbf{v}_i(t+1) = \omega(t)\mathbf{v}_i(t) + c_1r_1(pbest_i - \mathbf{x}_i) + c_2r_2(gbest - \mathbf{x}_i) \quad (15)$$

$$\mathbf{x}_i(t+1) = \mathbf{x}_i(t) + \mathbf{v}_i(t+1) \quad (16)$$

که در روابط بالا،  $c_1$  ضریب شخصی یادگیری است که برای بهترین موقعیت هر ذره لحاظ می‌شود.  $c_2$  ضریب یادگیری جمعی است که برای بهترین موقعیت تمامی ذرات لحاظ می‌گردد.  $r_1$  و  $r_2$  اعدادی هستند که به‌صورت تصادفی با توزیع یکنواخت در بازه ۰ تا ۱ انتخاب می‌شوند.  $pbest$  بهترین موقعیتی است که یک ذره تاکنون داشته است.  $gbest$  بهترین موقعیتی است که تا به حال توسط کل ذرات به‌دست آمده است.  $\omega$  پارامتر وزن لختی‌گر تکرار نامیده می‌شود. پارامتر  $\omega$  نقش مهمی را در همگرایی الگوریتم PSO ایفا می‌کند به طوری که مقادیر زیاد آن باعث افزایش مقدار بردار سرعت می‌شود و در هر تکرار گام‌های بزرگ‌تری برداشته خواهد شد و با کوچک شدن آن مقدار این گام‌ها کوچک‌تر می‌شود که برای رسیدن به جواب بهینه در گام‌های آخر مفید می‌باشد، از این رو با قرار دادن یک رابطه پویا به جای

<sup>1</sup> particle

<sup>2</sup> swarm

<sup>3</sup> inertia weight

مقداری ثابت برای  $\omega$  به این مهم می توان دست یافت. شی و ابرهارت<sup>۱</sup> یک رابطه خطی برای  $\omega$  پیشنهاد داده اند که به صورت رابطه زیر می باشد [۳۷]:

$$\omega = \omega_{\max} - \frac{\omega_{\max} - \omega_{\min}}{t_{\max}} t \quad (17)$$

که در آن،  $t_{\max}$  حداکثر تعداد تکرارها می باشد.  $\omega_{\min}$  و  $\omega_{\max}$  مقادیر حداکثر و حداقل  $\omega$  می باشد. یکی از ضعف های عمده الگوریتم PSO، گیر افتادن در بهینه محلی می باشد. برای رفع این نقیصه، اصلاح زیر در به هنگام سازی سرعت هر ذره بر مبنای تجمع غیرفعال پیشنهاد شده است [۳۸]:

$$v_i(t+1) = \omega(t)v_i(t) + c_1r_1(pb_{est_i} - x_i) + c_2r_2(gbest - x_i) + c_3r_3(R_i - x_i) \quad (18)$$

که در آن،  $R_i$  ذره ای تصادفی است که بصورت تصادفی از بین کل ذرات انتخاب می شود.  $r_3$  عددی که بصورت تصادفی با توزیع یکنواخت در بازه ۰ تا ۱ انتخاب می شوند.

### مدل ترکیبی LS-SVM و PSO

هنگام طراحی یک مدل موثر، مقادیر دو پارامتر اساسی در LS-SVM باید از قبل با دقت انتخاب شوند. این پارامترها شامل پارامترهای تنظیم  $\square$  و  $\sigma$  تابع کرنل GRBF می باشد. پارامتر تنظیم  $\square$  هزینه مبادله بین به حداقل رساندن خطای آموزشی و به حداقل رساندن پیچیدگی مدل را تعیین می کند. یک  $\square$  بزرگ جریمه های بالاتری را به خطاها اختصاص می دهد به طوری که رگرسیون آموزش داده می شود تا خطا را با تعمیم کمتر به حداقل برساند. در حالی که یک  $\square$  کوچک جریمه های کمتری را به خطاها اختصاص می دهد و این اجازه را می دهد تا حاشیه با خطا به حداقل برسد و بنابراین توانایی تعمیم بالاتر مدل LS-SVM را فراهم می کند [۳۹]. پارامتر  $\sigma$  تابع کرنل GRBF نگاشت غیرخطی را از فضای ورودی به برخی از فضای ویژگی های با ابعاد بالا تعریف می کند. انتخاب نامناسب پارامترها منجر به بیش برآزش یا کم برآزش داده ها می شود. به طور کلی، اکثر محققان هنوز از روش استاندارد (آزمایش و خطا) برای انتخاب پارامترها استفاده می کنند [۴۰]. ابتدا چند مدل LS-SVM بر اساس مجموعه پارامترهای مختلف تشکیل می شود، سپس آن ها را روی مجموعه اعتبارسنجی آزمایش کرده، تا پارامترهای بهینه را به دست آورند. با این حال، این روش زمان بر می باشد و ممکن است مقادیر بدست آمده برای پارامترها، مقادیر بهینه ای نباشند.

در این مطالعه، پارامترهای تنظیم  $\square$  و  $\sigma$  تابع کرنل GRBF با استفاده از الگوریتم PSO تعیین می شود. برای این منظور، می بایست یک تابع هدف مناسب برای مجموعه ای از داده های آموزش انتخاب شود اما باید از خطر بیش برآزش داده های آموزش در فرایند بهینه سازی جلوگیری شود. این مشکل را می توان با استفاده از تکنیک اعتبارسنجی متقاطع حل نمود [۴۰]. در این زمینه، اعتبارسنجی متقاطع k-fold بهترین تعادل را بین هزینه محاسباتی و برآورد پارامترهای قابل اعتماد ارائه می دهد [۴۱]. در این روش به ازای هر ذره در الگوریتم PSO، مدل LS-SVM با مجموعه ای از پارامترهای مرتبط به آن ذره ایجاد می شود و داده های آموزش به  $k$  زیرمجموعه تقسیم بندی می شود. سپس  $k-1$  زیرمجموعه به عنوان مجموعه آموزشی مدل LS-SVM استفاده می شود و عملکرد مدل LS-SVM با استفاده از معیار میانگین قدرمطلق درصد خطا (MAPE)<sup>۵</sup> برای  $k$  امین زیرمجموعه اندازه گیری می شود. روش فوق  $k$  بار تکرار می شود، به طوری

<sup>1</sup> Shi and Eberhart

<sup>2</sup> passive congregation

<sup>3</sup> over-fitting

<sup>4</sup> under-fitting

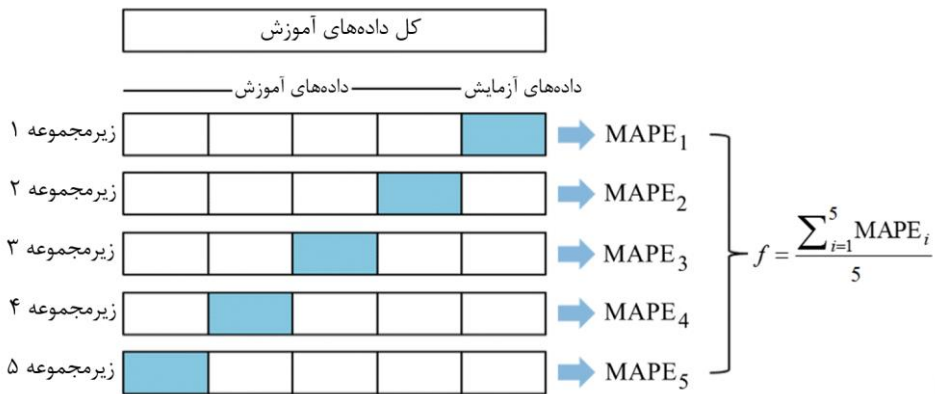
<sup>5</sup> Mean Absolute Percentage Error

که هر زیرمجموعه یک بار برای آزمایش استفاده می‌شود. بنابراین کمینه کردن میانگین MAPE در  $k$  آزمون تخمینی از خطا به عنوان تابع هدف هر ذره در نظر گرفته شده و به صورت زیر تعریف می‌شود:

$$\text{Minimize: } f = \frac{\sum_{i=1}^k \text{MAPE}_i}{k} \quad (19)$$

$$\text{MAPE}_i = \frac{\sum_{l=1}^l |a_i - p_i| / a_i}{l}, \quad i = 1, 2, \dots, k \quad (20)$$

که در روابط فوق،  $a_i$  و  $p_i$  به ترتیب مقدار واقعی و مقدار پیش‌بینی شده است.  $l$  معرف تعداد داده‌های آزمایش در  $k$ امین زیرمجموعه است. در این مطالعه، مقدار  $k$  در اعتبارسنجی متقاطع  $k$ -fold برابر ۵ فرض شده است. همچنین فرایند اعتبارسنجی متقاطع  $k$ -fold برای مقدار ۵ در شکل ۱ نمایش داده شده است.



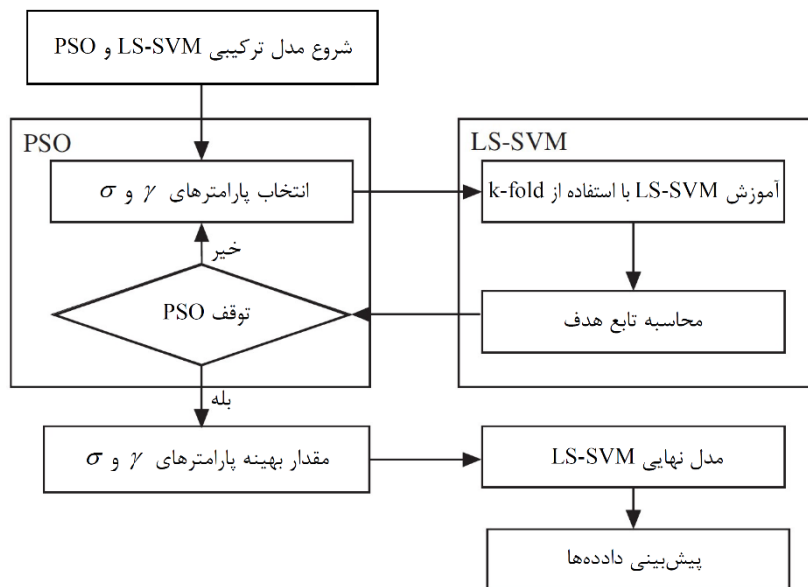
شکل ۱. فرایند اعتبارسنجی متقاطع  $k$ -fold برای مقدار ۵.

نهایتاً در شکل ۲ فلوجارت مدل ترکیبی LS-SVM و PSO نشان داده شده است.

### پایگاه داده‌های آزمایشگاهی IRI

در سال‌های اخیر، چندین مدل برای پیش‌بینی IRI توسط محققین پیشنهاد شده است. برخی از این مدل‌ها بر اساس پایگاه داده برنامه عملکرد بلندمدت روسازی (LTPP) استخراج شده‌اند، در حالی که برخی دیگر بر اساس اندازه‌گیری‌های مستقیم یا پایگاه‌های اطلاعاتی آژانس‌های محلی توسعه یافته‌اند. در این مطالعه، از پایگاه داده استاندارد LTPP گزارش شده در تحقیق آژی و لوب [۴۲] برای پیش‌بینی IRI استفاده شد. LTPP در ابتدا به عنوان یک برنامه ۲۰ ساله برای نظارت و جمع‌آوری داده‌ها در مورد انواع مختلف روسازی طراحی شد. این برنامه با هماهنگی اداره بزرگراه فدرال (FHWA)، با هدف بهبود و توسعه یک فرآیند طراحی شده برای روسازی‌های جدید و بازسازی شده، ارزیابی شرایط روسازی موجود، توسعه روش‌شناسی برای بهبود فرآیندهای طراحی و نگهداری موجود، و تعیین تأثیر متغیرهای

ساخت و ساز، شرایط محیطی، ویژگی های ترافیکی، و خواص مواد بر عملکرد سازه روسازی های انعطاف پذیر و بتنی انجام شد. برنامه LTPP از سال ۱۹۸۹ داده ها را از صدها بخش آزمایشگاهی در سراسر ایالات متحده و کانادا جمع آوری کرده است و پایگاه داده آن برای توسعه مدل های روسازی عالی تلقی می شود [۴۳].



شکل ۲. فلوچارت مدل ترکیبی PSO و LS-SVM

IRI تحت تأثیر عوامل متعددی از جمله بارهای ترافیکی، شرایط محیطی و اقلیمی، ناهمواری روسازی و کیفیت مصالح مورد استفاده در ساخت روسازی است. مدل پیش بینی گزارش شده در مطالعه بر اساس دو مجموعه پایگاه داده توسعه داده شده است. اولین مجموعه داده، که در جدول ۱ خلاصه شده است، شامل داده های IRI جمع آوری شده همراه با عدد ساختاری (SN)، سن (AGE)، و تعداد کل محور هم ارز عبوری (ESAL) است [۴۳]. مجموعه داده دوم شامل IRI اولیه اندازه گیری شده (IRI<sub>0</sub>)، سن اولیه (AGE<sub>0</sub>)، تعداد کل محور هم ارز عبوری اولیه (ESAL<sub>0</sub>)، SN، تفاوت در سن (ΔAGE)، تفاوت در تعداد کل محور هم ارز عبوری (ΔESAL) و IRI در سطح قابلیت اطمینان P (IRI<sub>P</sub>) می باشد [۴۳]. جدول ۲ آمار توصیفی متغیرهای این پایگاه را نشان می دهد.

جدول ۱. آمار توصیفی مجموعه اول داده های LTPP

متغیر	حداقل	حداکثر	میانگین	میانه	انحراف معیار
AGE (روز)	۸۸۷	۱۶۵۰۳	۴۴۱۲	۳۹۸۰	۳۰۵۲
ESAL (میلیون)	۰/۱۴۹	۱۹/۴۹۸	۳/۰۳۱	۰/۹۸	۵/۲۱
SN	۳/۲۳	۷/۲۲	۴/۹۲	۴/۶۱	۱/۱۷
IRI (m/km)	۰/۷۰۶	۲/۸۰۱	۱/۳۸۳	۱/۳۷	۰/۴۷

جدول ۲. آمار توصیفی مجموعه دوم داده‌های LTPP

متغیر	حداقل	حداکثر	میانگین	میانه	انحراف معیار
IRI <sub>0</sub> (m/km)	۰/۵۹۴	۲/۹۱۸	۱/۳۴۷	۱/۲۸	۰/۵۳۱
AGE <sub>0</sub> (روز)	۱۰۸۱	۱۶۵۰۳	۴۴۴۹	۴۰۴۷	۲۹۱۰
ESAL <sub>0</sub> (میلیون)	۰/۱۸۹	۱۹/۴۹۸	۲/۹۲۱	۱/۰۴	۴/۸۹۱
SN	۲/۸۵	۷/۲۲	۴/۷۵۳	۴/۵۷۲	۱/۱۷
ΔAGE (روز)	۰/۲۶۸	۰/۸۹	۱/۱۹۴	۰/۹۹	۰/۵۳۱
ΔESAL (میلیون)	۰/۰۷۷	۱۹/۲۳۳	۴/۷۲۴	۱/۶۳۹	۳/۵۲
IRI <sub>P</sub> (m/km)	۰/۵۹۴	۳/۱۴۳	۱/۴۲۶	۱/۳۵	۰/۵۸۳

## نتایج عددی

### نرمال‌سازی و تقسیم‌بندی پایگاه داده‌ها

یک مرحله مهم قبل از آماده‌سازی مدل ترکیبی LS-SVM و PSO، نرمال‌سازی پایگاه داده‌ها است. هدف از نرمال‌سازی این است که داده‌ها در یک محدوده مساوی تعریف شوند. در بسیاری از کاربردها، نایکسان بودن محدوده مقادیر ویژگی‌های (ورودی‌ها) مدل پیش‌بینی‌کننده باعث بی‌کیفیتی مدل می‌شود. برای رفع این مشکل، می‌توان از نرمال‌سازی داده‌ها استفاده کرد. نرمال‌سازی داده‌ها با یک تبدیل خطی یا غیرخطی، داده‌ها را به بازه‌ای که معمولاً [0, 1] است، انتقال می‌دهد. در این مطالعه، از رابطه زیر برای نرمال‌سازی داده‌های ورودی جداول (۱) و (۲) استفاده شده است [۴۴]:

$$\bar{x}_i = b_1 \frac{x_i - x_{\min,i}}{x_{\max,i} - x_{\min,i}} + b_2 \quad (21)$$

که در آن،  $x_{\min,i}$  و  $x_{\max,i}$  به ترتیب مقادیر نرمال شده، حداقل و حداکثر ورودی نام پایگاه داده می‌باشد. در این مطالعه،  $b_1$  و  $b_2$  به ترتیب برابر با ۰/۲ و ۰/۶ در نظر گرفته شد، به طوری که مقادیر نرمال شده بین ۰/۲ و ۰/۸ قرار می‌گیرند. هر یک از پایگاه داده‌ها شامل ۹۸ نمونه است که در گام بعدی، هر یک از پایگاه‌های داده‌ها به طور تصادفی به مجموعه‌های آموزشی و آزمایشی شامل ۶۹ (۷۰٪) و ۲۹ (۳۰٪) نمونه تقسیم شده‌اند.

### ارزیابی مدل ترکیبی PSO و LS-SVM

برای آماده‌سازی مدل ترکیبی PSO و LS-SVM، ابتدا پارامترهای تنظیم‌کننده الگوریتم PSO انتخاب شده است. به این منظور، ضرایب شخصی یادگیری و یادگیری جمعی به ترتیب برابر با ۰/۵، تعداد جمعیت (ذره) برابر ۵۰ و تعداد تکرار الگوریتم ۲۰۰ فرض شده است. همچنین، مقادیر حداکثر و حداقل  $\omega$  به ترتیب ۰/۹ و ۰/۴ انتخاب شده‌اند. باتوجه به ماهیت تصادفی الگوریتم PSO، این الگوریتم ۱۰ بار مستقل برای هر یک از داده‌ها اجرا شده و سپس کمترین مقدار تابع هدف در بین اجراهای مختلف به عنوان جواب بهینه انتخاب شده است. همچنین در اعتبارسنجی متقاطع k-fold، مقدار k برابر ۵ فرض شده است.

عملکرد مدل ترکیبی PSO و LS-SVM در پیش‌بینی IRI بر حسب سه معیار آماری شامل میانگین درصد مطلق خطا (MAPE)، ریشه میانگین مربعات خطا (RMSE) و ضریب تعیین ( $R^2$ ) ارزیابی شد. این معیارهای RMSE و  $R^2$  به صورت زیر بیان می‌شوند:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (a_i - p_i)^2}{n}} \quad (22)$$

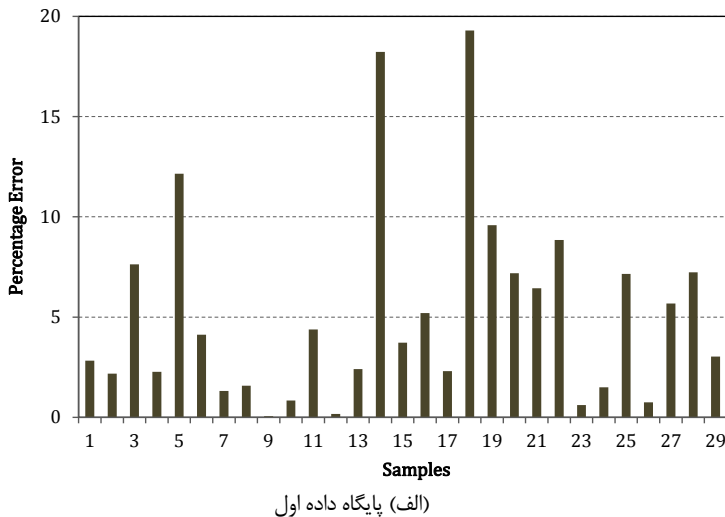
$$R^2 = \frac{\left( \sum_{i=1}^n (a_i - a_{ave})(p_i - p_{ave}) \right)^2}{\sum_{i=1}^n (a_i - a_{ave})^2 \sum_{i=1}^n (p_i - p_{ave})^2} \quad (23)$$

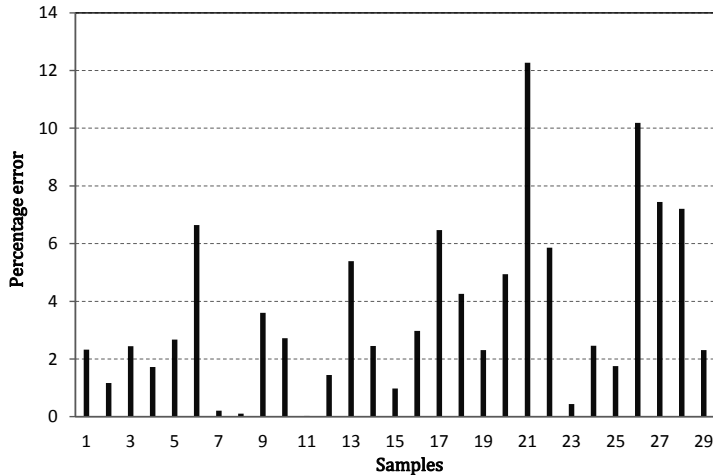
که در روابط بالا،  $a_{ave}$  و  $p_{ave}$  میانگین مقادیر واقعی و پیش‌بینی شده نمونه در پایگاه داده‌ها می‌باشند. جدول ۳ عملکرد پیش‌بینی مدل ترکیبی LS-SVM و PSO را برای مجموعه‌های آموزشی و آزمایشی دو پایگاه داده نشان می‌دهد. مقایسه داده‌ها نشان می‌دهد که مدل ترکیبی LS-SVM و PSO با دقت بسیار بالایی توانسته است IRI را در دو پایگاه داده پیش‌بینی نماید.

جدول ۳. عملکرد مدل ترکیبی PSO و LS-SVM در پیش‌بینی IRI

آزمایش			آموزش			پایگاه داده
$R^2$	RMSE	MAPE	$R^2$	RMSE	MAPE	
۰/۹۹۷	۰/۰۵۵	۵/۱۲۶	۰/۹۹۹	۰/۰۱۲	۱/۱۸۷	۱
۰/۹۹۴	۰/۰۷۴	۳/۹۱۷	۰/۹۹۶	۰/۰۵۷	۴/۳۳۴	۲

در ادامه، عملکرد مدل پیشنهادی با استفاده از درصد خطا بین مقادیر IRI مشاهده شده و پیش‌بینی شده برای پایگاه داده اول و دوم در شکل ۳ نشان داده شده است.

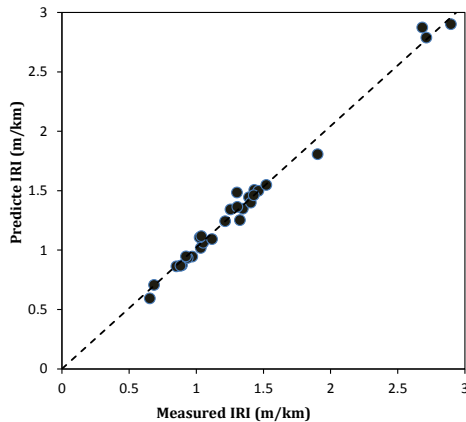




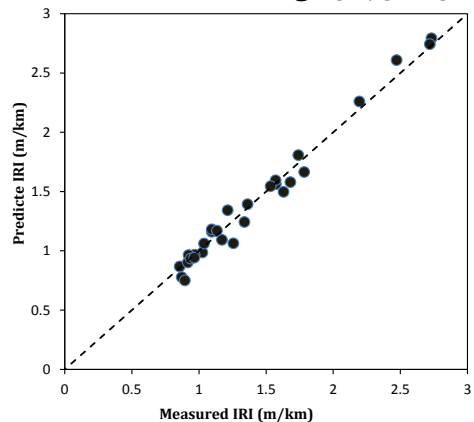
(الف) پایگاه داده دوم

شکل ۳. درصد خطای مدل پیشنهادی برای داده‌های آزمایش

نمودارهای پراکندگی بین مقادیر IRI اندازه‌گیری شده (واقعی) و پیش‌بینی شده در فرایند آزمایش برای هر دو پایگاه داده در شکل ۴ نشان داده شده است. نتایج در شکل ۴ نشان می‌دهد که مدل ترکیبی پیشنهادی به خوبی توانسته است مقدار IRI را پیش‌بینی نماید.



(ب) پایگاه داده دوم



(الف) پایگاه داده اول

شکل ۴. نمودارهای پراکندگی بین مقادیر IRI اندازه‌گیری شده و پیش‌بینی شده

### مقایسه عملکرد مدل ترکیبی پیشنهادی با دیگر روش‌ها

در این بخش، عملکرد مدل ترکیبی LS-SVM و PSO در مقایسه با دیگر روش‌ها برای پایگاه داده اول ارزیابی شده است. برای این منظور، در جدول ۴ مقایسه عملکرد مدل ترکیبی LS-SVM و PSO در این مطالعه با چند روش گزارش شده در مطالعات قبلی از جمله شبکه عصبی (ANN) [۴۲]، سیستم استنتاج عصبی فازی تطبیقی (ANFIS) [۴۵]، برنامه‌ریزی ژنتیک (GEP) [۴۳] و ترکیبی شبکه عصبی و برنامه‌ریزی ژنتیک (GEP-ANN) [۴۳] بر حسب ضریب همبستگی ( $R$ ) نشان داده شده است:

## جدول ۴. مقایسه عملکرد مدل ترکیبی LS-SVM و PSO با دیگر روش‌ها

ضریب همبستگی ( $R$ )	روش
۰/۹۲۵	ANN
۰/۹۸۶	ANFIS
۰/۹۰۵	GEP
۰/۹۹۴	GEP-ANN
۰/۹۹۸	LS-SVM and PSO

نتایج جدول (۴) نشان می‌دهد که مدل پیشنهادی LS-SVM و PSO در این مطالعه از دقت بالایی در پیش‌بینی نسبت به دیگر روش‌ها برخوردار می‌باشد. لذا با اطمینان از این مدل ترکیبی می‌توان در ارزیابی عملکرد روسازی استفاده نمود.

## نتیجه‌گیری

در این مطالعه، مدل ترکیبی LS-SVM و PSO برای پیش‌بینی IRI پیشنهاد شده است. الگوریتم PSO به‌منظور تعیین بهینه پارامترهای تنظیم‌کننده مدل LS-SVM استفاده شده است. همچنین، تکنیک اعتبارسنجی متقاطع  $k$ -fold برای جلوگیری از بیش‌برازش مدل LS-SVM در پروسه آموزش در نظر گرفته شده است. دو پایگاه داده آزمایشگاهی برنامه عملکرد بلندمدت روسازی (LTPP) جهت پیش‌بینی IRI برای سنجش کارایی مدل ترکیبی پیشنهادی استفاده شده است. نتایج ارزیابی دقت مدل ترکیبی LS-SVM و PSO نشان داد که معیار ضریب تعیین ( $R^2$ ) برای دو پایگاه داده LTPP برابر ۰/۹۹۴ و ۰/۹۹۷ به‌دست‌آمده است که نشان‌دهنده دقت بالای این مدل پیشنهادی در پیش‌بینی IRI می‌باشد. در پایان، مقایسه عملکرد مدل ترکیبی با دیگر روش‌های یادگیری ماشین بر حسب ضریب همبستگی ( $R$ ) انجام شد. نتایج نشان داد که مدل ترکیبی LS-SVM و PSO از دقت بالایی برخوردار است؛ بنابراین، مدل پیشنهادی به‌طور مؤثری می‌تواند در پیش‌بینی عملکرد روسازی مورد استفاده قرار گیرد.

## References

- [1] Tamagusko, T., & Ferreira, A. (2023). *Machine Learning for Prediction of the International Roughness Index on Flexible Pavements: A Review, Challenges, and Future Directions*. *Infrastructures*, 8(12), 170. <https://doi.org/10.3390/infrastructures8120170>.
- [2] Kulkarni, R. B., & Miller, R. W. (2003). *Pavement Management Systems: Past, Present, and Future*. *Transportation Research Record*, 1853(1), 65-71. <https://doi.org/10.3141/1853-08>.
- [3] Gupta, A., Kumar, P., & Rastogi, R. (2014). *Critical review of flexible pavement performance models*. *KSCE Journal of Civil Engineering*, 18(1), 142-148. <https://doi.org/10.1007/s12205-014-0255-2>.
- [4] Santos, J., & Ferreira, A. (2012). *Pavement design optimization considering costs and preventive interventions*. *Journal of Transportation Engineering*, 138(7), 911-923. [https://doi.org/10.1061/\(ASCE\)TE.1943-5436.0000390](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000390).
- [5] Huang, M.-H., & Rust, R. T. (2018). *Artificial intelligence in service*. *Journal of service research*, 21(2), 155-172. <https://doi.org/10.1177/1094670517752459>.
- [6] Wang, W., & Siau, K. (2019). *Artificial intelligence, machine learning, automation, robotics, future of work and future of humanity: A review and research agenda*. *Journal of Database Management (JDM)*, 30(1), 61-79. <https://doi.org/10.4018/JDM.2019010104>.

- [7] Gong, H., Sun, Y., Hu, W., Polaczyk, P. A., & Huang, B. (2019). *Investigating impacts of asphalt mixture properties on pavement performance using LTPP data through random forests*. *Construction and Building Materials*, 204, 203-212. <https://doi.org/10.1016/j.conbuildmat.2019.01.198>.
- [8] Jarrahi, H., Asadi, A., Khatibinia, M., & Etedali, S. (2020). *Optimal design of rotational friction dampers for improving seismic performance of inelastic structures*. *Journal of Building Engineering*, 27, 100960. <https://doi.org/10.1016/j.jobbe.2019.100960>.
- [9] Khatibinia, M., Akbari, S., Yazdani, H., & Gharehbaghi, S. (2023). *Damage-based optimal control of steel moment-resisting frames equipped with tuned mass dampers*. *Journal of Vibration and Control*, 30(3-4), 659-672. <https://doi.org/10.1177/10775463221149462>.
- [10] Jordan, M. I., & Mitchell, T. M. (2015). *Machine learning: Trends, perspectives, and prospects*. *Science*, 349(6245), 255-260. <https://doi.org/10.1126/science.aaa8415>.
- [11] Abduljabbar, R., Dia, H., Liyanage, S., & Bagloee, S. A. (2019). *Applications of Artificial Intelligence in Transport: An Overview*. *Sustainability*, 11(1), 189. <https://doi.org/10.3390/su11010189>.
- [12] Mehdipour, M., Chitsaz, E., & Etemadi, M. (2024). *A Perspective on Human Interaction and Artificial Intelligence: Bibliometric Analysis with Co-occurrence Technique*. *Karafan Journal*, 21(3), 13-33. <https://doi.org/10.48301/kssa.2024.428257.2778>.
- [13] Deypir, M., Maarefi, A., & Zoughi, T. (2024). *A Hybrid Multivariate Routing Approach for Improving Efficiency in VANETS*. *Karafan Journal*, 21(1), 39-62. <https://doi.org/10.48301/kssa.2024.423334.2752>.
- [14] Rahimi, K., Agha gholizadeh sayar, A., & Izadyar, M. (2024). *Presenting a Sustainable Performance Strategy in the Automotive Supply Chain Using Fuzzy Network Analysis*. *Karafan Journal*, 21(1), 335-354. <https://doi.org/10.48301/kssa.2023.414800.2691>.
- [15] Géron, A. (2022). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*. O'Reilly Media, Inc.
- [16] Yao, L., Leng, Z., Jiang, J., & Ni, F. (2021). *Modelling of pavement performance evolution considering uncertainty and interpretability: a machine learning based framework*. *International Journal of Pavement Engineering*, 23, 5211 - 5226. <https://doi.org/10.1080/10298436.2021.2001814>.
- [17] Justo-Silva, R., Ferreira, A., & Flintsch, G. (2021). *Review on machine learning techniques for developing pavement performance prediction models*. *Sustainability*, 13(9), 5248. <https://doi.org/10.3390/su13095248>.
- [18] Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., & Brennan, S. E. (2021). *The PRISMA 2020 statement: an updated guideline for reporting systematic reviews*. *bmj*, 372. <https://doi.org/10.1136/bmj.n71>.
- [19] Sayers, M. W., Gillespie, T. D., & Queiroz, C. A. (1986). *The international road roughness experiment: Establishing correlation and a calibration standard for measurements*. <http://documents.worldbank.org/curated/en/326081468740204115/The-International-Road-Roughness-Experiment-IRRE-establishing-correlation-and-a-calibration-standard-for-measurements>.
- [20] Hoque, Z. (2006). *Highway condition surveys and serviceability evaluation*.

- [21] Hall, K., & Muñoz, C. (1999). *Estimation of present serviceability index from international roughness index*. Transportation Research Record, 1655(1), 93-99. <https://doi.org/10.3141/1655-13>.
- [22] Hossain, M. I., Gopiseti, L. S. P., & Miah, M. S. (2019). *International Roughness Index Prediction of Flexible Pavements Using Neural Networks*. Journal of Transportation Engineering, Part B: Pavements, 145(1), 04018058. <https://doi.org/10.1061/JPEODX.0000088>.
- [23] Abdelaziz, N., Abd El-Hakim, R. T., El-Badawy, S. M., & Afify, H. A. (2020). *International Roughness Index prediction model for flexible pavements*. International Journal of Pavement Engineering, 21(1), 88-99. <https://doi.org/10.1080/10298436.2018.1441414>.
- [24] Zeiada, W., Dabous, S. A., Hamad, K., Al-Ruzouq, R., & Khalil, M. A. (2020). *Machine learning for pavement performance modelling in warm climate regions*. Arabian journal for science and engineering, 45(5), 4091-4109. <https://doi.org/10.1007/s13369-020-04398-6>.
- [25] Gong, H., Sun, Y., Shu, X., & Huang, B. (2018). *Use of random forests regression for predicting IRI of asphalt pavements*. Construction and Building Materials, 189, 890-897. <https://doi.org/10.1016/j.conbuildmat.2018.09.017>.
- [26] Naseri, H., Jahanbakhsh, H., Foomajd, A., Galustanian, N., Karimi, M. M., & D. Waygood, E. (2023). *A newly developed hybrid method on pavement maintenance and rehabilitation optimization applying Whale Optimization Algorithm and random forest regression*. International Journal of Pavement Engineering, 24(2), 2147672. <https://doi.org/10.1080/10298436.2022.2147672>.
- [27] Sharma, A., Sachdeva, S. N., & Aggarwal, P. (2023). *Predicting IRI using machine learning techniques*. International Journal of Pavement Research and Technology, 16(1), 128-137. <https://doi.org/10.1007/s42947-021-00119-w>.
- [28] Naseri, H., Shokoohi, M., Jahanbakhsh, H., Karimi, M. M., & Waygood, E. (2023). *Novel soft-computing approach to better predict flexible pavement roughness*. Transportation Research Record, 2677(10), 246-259. <https://doi.org/10.1177/036119812311610>.
- [29] Zhang, M., Gong, H., Jia, X., Xiao, R., Jiang, X., Ma, Y., & Huang, B. (2020). *Analysis of critical factors to asphalt overlay performance using gradient boosted models*. Construction and Building Materials, 262, 120083. <https://doi.org/10.1016/j.conbuildmat.2020.120083>.
- [30] Damirchilo, F., Hosseini, A., Mellat Parast, M., & Fini, E. H. (2021). *Machine learning approach to predict international roughness index using long-term pavement performance data*. Journal of Transportation Engineering, Part B: Pavements, 147(4), 04021058. <https://doi.org/10.1061/JPEODX.0000312>.
- [31] Song, Y., Wang, Y. D., Hu, X., & Liu, J. (2022). *An efficient and explainable ensemble learning model for asphalt pavement condition prediction based on LTPP dataset*. IEEE Transactions on Intelligent Transportation Systems, 23(11), 22084-22093. <https://doi.org/10.1109/TITS.2022.3164596>.
- [32] Cortes, C., & Vapnik, V. (1995). *Support-vector networks*. Machine Learning, 20(3), 273-297. <https://doi.org/10.1007/BF00994018>.
- [33] Suykens, J. A., & Vandewalle, J. (1999). *Least squares support vector machine classifiers*. Neural processing letters, 9(3), 293-300. <https://doi.org/10.1023/A:1018628609742>

- [34] Yu, L., Yao, X., Wang, S., & Lai, K. K. (2011). *Credit risk evaluation using a weighted least squares SVM classifier with design of experiment for parameter selection*. *Expert Systems with Applications*, 38(12), 15392-15399. <https://doi.org/10.1016/j.eswa.2011.06.023>.
- [35] Kennedy, J., & Eberhart, R. (1995, 27 Nov.-1 Dec. 1995). Particle swarm optimization. *Proceedings of ICNN'95 - International Conference on Neural Networks*.
- [36] Shayesteh Bilondi, M. R., Yazdani, H., & Khatibinia, M. (2018). *Seismic energy dissipation-based optimum design of tuned mass dampers*. *Structural and Multidisciplinary Optimization*, 58(6), 2517-2531. <https://doi.org/10.1007/s00158-018-2033-0>.
- [37] Shi, Y., & Eberhart, R. (1998, 4-9 May 1998). A modified particle swarm optimizer. 1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence (Cat. No.98TH8360).
- [38] He, S., Wu, Q. H., Wen, J. Y., Saunders, J. R., & Paton, R. C. (2004). *A particle swarm optimizer with passive congregation*. *Biosystems*, 78(1), 135-147. <https://doi.org/10.1016/j.biosystems.2004.08.003>.
- [39] Guo, X. C., Yang, J. H., Wu, C. G., Wang, C. Y., & Liang, Y. C. (2008). *A novel LS-SVMs hyper-parameter selection based on particle swarm optimization*. *Neurocomputing*, 71(16), 3211-3215. <https://doi.org/10.1016/j.neucom.2008.04.027>.
- [40] Chen, K.-Y. (2007). *Forecasting systems reliability based on support vector regression with genetic algorithms*. *Reliability Engineering & System Safety*, 92(4), 423-432. <https://doi.org/10.1016/j.ress.2005.12.014>.
- [41] Duan, K., Keerthi, S. S., & Poo, A. N. (2003). *Evaluation of simple performance measures for tuning SVM hyperparameters*. *Neurocomputing*, 51, 41-59. [https://doi.org/10.1016/S0925-2312\(02\)00601-X](https://doi.org/10.1016/S0925-2312(02)00601-X).
- [42] Ozbay, K., & Laub, R. (2001). *Models for pavement deterioration using LTPP*.
- [43] Mazari, M., & Rodriguez, D. D. (2016). *Prediction of pavement roughness using a hybrid gene expression programming-neural network technique*. *Journal of Traffic and Transportation Engineering (English Edition)*, 3(5), 448-455. <https://doi.org/10.1016/j.jtte.2016.09.007>.
- [44] Mohammadzadeh, M. R., Esfandnia, F & , Khatibinia, M. (2023). *Prediction of Shear Strength of Reinforced Concrete Deep Beams Using Neuro-Fuzzy Inference System and Meta-Heuristic Algorithms*. *Civil Engineering Infrastructures Journal*, 56(1), 137-157. <https://doi.org/10.22059/cej.2022.334953.1803>.
- [45] Terzi, S. (2013). *Modeling for pavement roughness using the ANFIS approach*. *Advances in Engineering Software*, 57, 59-64. <https://doi.org/10.1016/j.advengsoft.2012.11.013>.