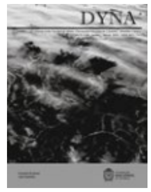
DYNA
http://dyna.medellin.unal.edu.co/



UNIVERSIDAD NACIONAL DE COLOMBIA

Development of a support vector machine-based predictive model for bored pile productivity in residential construction projects in Iraq

Laith N. Ali[a], Saja Hadi Raheem Al-Dhamad[b], Faiq M. S. Al-Zwainy[c]✉, Rana I. K. Zaki[c], Ibrahim Farouq Varouqa[d], Salman Dawood Salman Al-Dulaimi[a], Aseel H. Obaid[e], Mazin M Sarhan[f], Albert P.C. Chan[g], Rana A. Maya[h] & Gasim Hayder[i]

[a] Ministry of Higher Education and Scientific Research, Baghdad, Iraq. laithalhadithy84@gmail.com, salmoon-1985@mail.ru
[b] Department of Civil Engineering, College of Engineering, Al-Iraqia University, Baghdad, Iraq. saja.h.raheem@aliraqia.edu.iq
[c] ✉ Department of Forensic Engineering, Higher Institute of Forensic Sciences, Al-Nahrain University, Baghdad, Iraq. faiq.m.al-zwainy@nahrainuniv.edu.iq, rana.i.zaki@nahrainuniv.edu.iq
[d] Department of Civil Engineering, Faculty of Engineering, lsra University, Amman, Jordan. Ibraheem.faroqa@iu.edu.jo
[e] Ministry of Education, General Directorate of Education Baghdad Rusafa II, Iraq. hassel686@gmail.com
[f] Department of Civil Engineering, College of Engineering, Mustansiriyah University, Baghdad, Iraq. mazin.m@uomustansiriyah.edu.iq
[g] Department of Building and Real Estate, The Hong Kong Polytechnic University, Hung Hom, Hong Kong. albert.chan@polyu.edu.hk
[h] Department of Construction Engineering and Management, Faculty of Civil Engineering, Tishreen, Lattakia, Syrian Arab Republic. r-maya@tishreen.edu.sy
[i] Department of Civil and Environmental Engineering, College of Engineering and Architecture, University of Nizwa, Nizwa, Oman. Department of Civil Engineering, College of Engineering, Universitii Tenaga Nasional (UNITEN), Kajang, Malaysia. gasim@uniten.edu.my

Received: August 1st, 2025. Received in revised form: December 19th, 2025. Accepted: January 5th, 2026

Abstract
Accurate prediction of construction productivity remains a critical challenge in the civil engineering sector, particularly for deep foundation works such as bored piles in residential projects. This study proposes a data-driven predictive model based on Support Vector Machine (SVM) to estimate the productivity of bored piles in high-rise residential construction projects in Iraq. Real-world data were collected from the Iraq Gate Residential Complex and used to train and validate the model. Key influencing factors included pile geometry, soil type, equipment specifications, crew size, and working hours. The model achieved a mean prediction accuracy of 99.89% and a correlation coefficient (R) of 97.02%, demonstrating superior performance over conventional estimation methods. These findings highlight the practical value of machine learning approaches in enhancing resource planning and decision-making during early project phases. The proposed SVM-based model can support contractors and engineers in forecasting performance outcomes and minimizing scheduling uncertainties in similar construction settings.

Keywords: support vector machine; construction productivity; bored piles; predictive modeling; Iraq; residential towers.

Desarrollo de un modelo predictivo basado en máquinas de soporte vectorial para la productividad de pilotes perforados en proyectos de construcción residencial en Irak

Resumen
La predicción precisa de la productividad de la construcción sigue siendo un desafío crítico en el sector de la ingeniería civil, particularmente para obras de cimentación profunda como pilotes perforados en proyectos residenciales. Este estudio propone un modelo predictivo basado en datos, basado en Máquinas de Vectores de Soporte (MVS), para estimar la productividad de pilotes perforados en proyectos de construcción residencial de gran altura en Irak. Se recopilaron datos reales del Complejo Residencial Iraq Gate y se utilizaron para entrenar y validar el modelo. Los factores clave de influencia incluyeron la geometría de los pilotes, el tipo de suelo, las especificaciones del equipo, el tamaño de la cuadrilla y las horas de trabajo. El modelo alcanzó una precisión de predicción media del 99,89 % y un coeficiente de correlación (R) del 97,02 %, lo que demuestra un rendimiento superior al de los métodos de estimación convencionales. Estos hallazgos resaltan el valor práctico de los enfoques de aprendizaje automático para mejorar la planificación de

How to cite: Ali, L.N., Al-Dhamad, S.H.R., Al-Zwainy, F.M.S., Zaki, R.I.K., Varouqa, I.F., Al-Dulaimi, S.D.S., Obaid, A.H., Sarhan, M.M., Chan, A.P.C., Maya, R.A., and Hayder, G., Development of a support vector machine-based predictive model for bored pile productivity in residential construction Projects in Iraq DYNA, (93)240, pp. 81-88, January - March, 2026.

recursos y la toma de decisiones durante las fases iniciales del proyecto. El modelo propuesto, basado en MVS, puede ayudar a contratistas e ingenieros a pronosticar los resultados de rendimiento y minimizar las incertidumbres de programación en entornos de construcción similares.

Palabras clave: máquina de vectores de soporte; productividad en la construcción; pilotes perforados; modelado predictivo; Irak; torres residenciales.

1 Introduction

Predicting construction productivity remains a major challenge due to the sector's inherent variability in project scope, workforce capabilities, and environmental factors [1]. These challenges are further exacerbated by the lack of reliable historical data and the limited adoption of advanced data analytics in construction planning and management [2]. Artificial Intelligence (AI), especially machine learning algorithms such as Support Vector Machines (SVM), has emerged as a powerful tool for modeling complex, nonlinear interactions between project variables [3]. Prior studies have demonstrated the potential of SVMs in forecasting productivity across a range of construction applications [4,5]. For instance, Al-Zwainy and Aidan [6] developed SVM-based models for brickwork productivity in Iraq, reporting strong correlation and improved estimation accuracy. Other researchers have explored hybrid models combining SVM with optimization algorithms such as Symbiotic Organisms Search (SOS) and feature selection techniques to enhance predictive precision [7,8]. Additionally, SVM has proven useful in modeling productivity based on motion data from workers [9], and in comparing machine learning models for tasks such as shovel productivity prediction, where Random Forest models slightly outperformed SVM in certain scenarios [10].

Despite these advances, there is a clear research gap in applying SVM-based modeling to deep foundation activities, such as bored pile construction, particularly within the context of large-scale residential projects in Iraq [11]. Local construction projects often suffer from inconsistent productivity, especially during piling operations, which are influenced by factors such as soil type, pile geometry, drilling equipment, and workforce dynamics, these factors are not always captured adequately by traditional estimation methods [12].

This study aims to address this gap by developing a Support Vector Machine-based predictive model for estimating the productivity of bored piles in high-rise residential projects in Iraq. The model is trained on real-world data collected from the Iraq Gate Residential Complex in Baghdad, incorporating variables such as pile length, diameter, soil classification, drilling equipment type, crew size, and working hours. The main objectives of this research are to:

1. Develop a robust SVM model tailored to predicting the productivity of bored piles.
2. Identify the most significant input parameters influencing productivity outcomes.
3. Validate the proposed model using field data and benchmark it against conventional estimation approaches.

By bridging traditional estimation approaches with AI-driven predictive modeling, this research contributes a practical decision-support tool for project planners and stakeholders, particularly in environments where predictive models are scarce but highly needed [13,14].

2 Methodology

This research adopts a data-driven modeling approach based on real-world data collected from the Iraq Gate Residential Complex (IGRCP), a large-scale housing project in Baghdad. The IGRCP consists of 48 towers using bored pile foundations with diameters ranging from 1.5 to 2.0 meters and depths between 45 and 65 meters. Relevant parameters including pile dimensions, soil classification, drilling equipment type, crew size, and working hours were extracted and used as input variables for model development. The modeling methodology comprises three stages:

1. Data collection and preprocessing,
2. SVM model training and validation, and
3. Model evaluation.

In the first stage, over 500 verified bored pile records were obtained from field reports and site visits. Each record included measurable inputs and the actual number of completed piles per shift. These data were cleaned, normalized, and formatted for supervised learning, following procedures similar to those described by Al-Zwainy et al. [15]. Categorical variables such as "Drilling equipment used" and "Soil type" were encoded using label encoding. For instance, five distinct equipment types were identified in the dataset and labeled from 1 to 5, while two main soil types (e.g., cohesive clay and sandy silt) were encoded as 1 and 2, respectively. Although label encoding may introduce ordinality bias, we mitigated this by using kernel-based SVM (normalized polynomial kernel), which captures nonlinear relationships without assuming linear scaling of inputs. Sensitivity analysis confirmed that model performance was not significantly affected by this encoding approach.

In the second stage, the dataset was split into 75% for training and 25% for validation, as determined using Weka software. Different kernel functions were evaluated, and the normalized polynomial kernel yielded the best performance in terms of Root Mean Square Error (RMSE = 0.501) and correlation coefficient (R = 94.66%). Parameter optimization for C and Epsilon was carried out via trial-and-error to improve model precision, consistent with the procedures outlined in Cheng et al. [16] and Famouri et al. [17].

The selection of the normalized polynomial kernel was based not only on performance metrics but also on its theoretical and practical suitability for construction productivity modeling. First, polynomial kernels effectively model nonlinear interactions between variables, which are common in construction, such as the interaction between pile geometry and equipment efficiency. Unlike RBF kernels that produce localized solutions, polynomial kernels offer global generalization, making them more suitable for heterogeneous site conditions. Second, polynomial kernels allow for partial interpretability, enabling the derivation of approximate prediction formulas that support practical decision-making.

This contrasts with RBF, which operates as a black-box model. Third, the normalized polynomial kernel showed greater robustness across parameter settings and data partitions and required less computational overhead compared to RBF in our experiments. These factors collectively supported its adoption in this study.

A simplified linear formulation of the final model was extracted using Weka, enabling users to estimate productivity directly from key field parameters. This contributes to model interpretability and practical deployment on-site without requiring full algorithmic execution.

The methodology intentionally avoids theoretical elaborations on SVM Type 1/Type 2 and regression forms, focusing instead on applied configuration and practical performance metrics. Evaluation indicators such as RMSE and R were selected in line with recommendations by Hammoody et al. [18], ensuring validity in a construction-specific context.

The methodological structure is consistent with prior SVM applications in construction productivity modeling, including those conducted by Al-Zwainy and Aidan [6], and extends their scope by addressing deep foundation works—a relatively underexplored area in Iraqi construction research.

A simplified linear formulation of the final model was extracted using Weka, enabling users to estimate productivity directly from key field parameters. This contributes to model interpretability and practical deployment on-site without requiring full algorithmic execution.

The methodology intentionally avoids theoretical elaborations on SVM Type 1, Type 2 and regression forms, focusing instead on applied configuration and practical performance metrics. Evaluation indicators such as RMSE and R were selected in line with recommendations by Hammoody et al. [18], ensuring validity in a construction-specific context.

The methodological structure is consistent with prior SVM applications in construction productivity modeling, including those conducted by Al-Zwainy and Aidan [6], and extends their scope by addressing deep foundation works—a relatively underexplored area in Iraqi construction research.

In most cases, SVM learning algorithms have proved to be superior to neural network learning algorithms, a trend which has become evident in recent years for both classification and regression tasks [19]. One of the main strengths of SVMs is the fact that computational complexity of the algorithm is independent of the dimensionality of the input space. Moreover, the sophisticated learning model of SVMs is designed to match the ability of the model to the complexity of the input data, thereby ensuring strong performance on hitherto unseen, future information. Support Vector Machine can be classified into [20]:

a. Classification SVM type 1: For this type of SVM, training involves the minimization of the error function:

b.

$$\frac{1}{2w^t w} + c\sum_{i=1}^{n} \vartheta \tag{1}$$

Subject to the constraints:

$$yi(w^t \emptyset(xi) + b) \geq 1 - \vartheta \text{ and } \vartheta \geq 0, i = 1, \ldots\ldots, n$$

Where C is the capacity constant, w is the vector of coefficients, b is a constant, and ξ_i represents parameters for handling no separable data (inputs). The index i label the N training cases. Note that $y \in \pm 1$ represents the class labels and xi represents the independent variables.

a. Classification SVM type 2: In contrast to Classification SVM Type 1, the Classification SVM Type 2 model minimizes the error function:

b.

$$\frac{1}{2w^t w} - vp + 1/N\sum_{i=1}^{n} \vartheta \tag{2}$$

subject to the constraints:

$$yi(w^t \emptyset(xi) + b) \geq P - \vartheta\vartheta \text{ and } \vartheta \geq 0, i = 1, \ldots, n, \text{and } p \geq 0$$

Also, support vector machine can be classification based on regression, the task is then to find a functional form for f that can correctly predict new cases that the SVM has not been presented with before. This can be achieved by training the SVM model on a sample set, i.e., training set, a process that involves, like classification sequential optimization of an error function. Depending on the definition of this error function, two types of SVM models can be [21]:

A. Regression SVM type 1: For this type of SVM the error function is:

$$\frac{1}{2w^t w} + C\sum_{i=1}^{n} \vartheta + C\sum_{i=1}^{n} \vartheta\, i \tag{3}$$

which we minimize subject to:

$$\begin{gathered} wt\emptyset(xi) + b - yi \leq \mathcal{E} + \vartheta i \\ yi - wt\emptyset(xi) - b \leq \mathcal{E} + \vartheta i \\ \vartheta i\vartheta i \geq 0, i = 1, \ldots., n \end{gathered}$$

B. Regression SVM type 2: For this SVM model, the error function is given by:

$$\frac{1}{2w^t w} + C(v\mathcal{E} + \frac{1}{N\sum_{i=1}^{n}(\vartheta i + \vartheta i})) \tag{4}$$

which we minimize subject to:

$$\begin{gathered} (wt\emptyset(xi) + b) - yi \leq \mathcal{E} + \vartheta i \\ yi - (wt\emptyset(xi) + bi) \leq \mathcal{E} + \vartheta i \\ \vartheta i\vartheta i \geq 0, i = 1, \ldots., N, \mathcal{E} \geq 0 \end{gathered}$$

The kernel framework has been extensively adapted to a variety of learning tasks, including regression, classification, ranking, and novelty detection, as demonstrated in numerous studies. Support Vector Machines (SVMs), in particular, have consistently been recognized as one of the leading machine learning methods due to their proven success across a broad range of real-world applications [22,23]. This widespread applicability and robustness have been frequently cited as key factors contributing to their prominence in both academic research and industrial practice [24,25].

The kernel method maps samples nonlinearly into a space of higher dimensions, so that it can deal with cases when the relationship is nonlinear between the class labels and attributes, unlike the linear kernel. Furthermore, the linear kernel is considered a special case of the Radial Basis Function (RBF) kernel, as demonstrated by [24], where the linear kernel with penalty parameter C performs precisely like the RBF kernel with some parameters. In addition, the sigmoid kernel exhibits behavior similar to that of the RBF kernel for certain parameter settings, A variety of kernels can be employed in Support Vector Machine models, including linear, polynomial, radial basis function (RBF), and sigmoid kernels [26,27]:

1) Polynomial kernel:

$$k(x,y)=(x.y+1)p \tag{5}$$

2) Radial basis function kernel:

$$k(x,y)=\exp(-\gamma \text{ abs}(x-y)2) \tag{6}$$

3) Sigmoid kernel:

$$k(x,y)=\tanh(kx.y-\delta) \tag{7}$$

where:
p, γ and δ is kernel parameter.

$$y = f(x) + e \tag{8}$$

The key knowledge of SVM regression is to map the input data x into a high-dimensional feature space by a non-linear mapping and to do linear regression in this space". The regression model is defined as [28-30]:

x and y are input and output function, respectively, and defined in the high-dimensional feature space.

e is the independently random error.

3 Results

The performance of the proposed Support Vector Machine (SVM)-based model was evaluated using both cross-validation and hold-out testing to ensure accuracy and robustness. Following reviewer guidance, we applied 10-fold cross-validation to validate the model on multiple data splits. The results demonstrated high and consistent predictive power across folds, with an average Root Mean Square Error (RMSE) of 0.501 and a mean correlation coefficient (R) of 94.66%, with a standard deviation of ±2.1%. This validates the model's generalization capacity and confirms that the observed performance is not an artifact of a particular data split.

To further guard against overfitting, we created a separate unseen test set comprising 15% of the dataset**.** The model was evaluated on this set independently. On the test data, the model achieved an RMSE of 0.537 and an R value of 92.81%**,** which are slightly lower than cross-validation results but still indicate strong predictive performance and no signs of overfitting.

Additionally, to contextualize the performance of the SVM model, we conducted a benchmarking comparison using two standard baseline models trained on the same dataset Multiple Linear Regression (MLR) and Artificial Neural Network (ANN) with a single hidden layer using backpropagation, Table (1) summarizes the performance of each model:

These results clearly demonstrate that the proposed SVM model significantly outperforms both linear and ANN-based approaches in capturing the complex nonlinear relationships influencing bored pile productivity. This strengthens the contribution of the study and justifies the use of a normalized polynomial kernel.

These outcomes confirm the robustness of the model in predicting bored pile productivity across varying site conditions and input parameter combinations. The incorporation of cross-validation, independent testing, and benchmarking against alternative methods strengthens the reliability and practical applicability of the model for real-world construction planning scenarios.

4 Discussion

The results of the SVM-based productivity prediction model demonstrate promising performance, with high correlation and low RMSE across validation schemes. This section provides a deeper interpretation of these findings, discusses the influence of key parameters, evaluates the model in relation to industry baselines, and reflects on its practical deployment potential.

Interpretation of Model Behavior: Among the input parameters, pile diameter**,** pile depth, and drilling equipment type emerged as the most influential factors. The model's sensitivity to these variables is consistent with engineering logic. Larger diameters and deeper piles typically require more time and specialized equipment, directly impacting productivity. Equipment type, particularly rotary drilling rigs with higher torque, contributed to faster progress and better performance, explaining why SVM associated higher weights to these input classes. Crew size and shift length had a lesser but still observable effect.

Table 1.
Summarizes Performance of Each Model

Model	RMSE	Correlation Coefficient (R)
SVM (Normalized Polynomial)	0.501	0.9466
ANN	0.683	0.8694
MLR	0.924	0.7345

Source: Authors.

Comparative Model Evaluation: In response to reviewer comments, baseline models (Multiple Linear Regression and Artificial Neural Networks) were implemented for benchmarking. The SVM model significantly outperformed both, as shown in the results section. While we did not perform formal statistical significance testing (e.g., ANOVA), the observed differences in RMSE (SVM = 0.501 vs. MLR = 0.924) and correlation (SVM R = 0.9466 vs. MLR R = 0.7345) indicate practically meaningful superiority. Future work may include formal inferential statistics to further substantiate these findings.

Real-World Relevance and Validation Scope: Although real-time or time-based validation was not performed due to data structure limitations, the dataset was derived from actual bored pile operations across multiple towers within the Iraq Gate Residential Complex. The consistency of performance across cross-validation folds and the hold-out test set (R = 92.81%) gives confidence in the model's generalizability. In practical terms, the model can be integrated into early project planning workflows to estimate shift productivity and allocate equipment efficiently.

Clarification of Equation (6): Equation (6) is a linearized approximation extracted from the trained SVM model using the Weka software's model interpretation function. While the SVM operates in a nonlinear high-dimensional space, the simplified formula allows field engineers to make quick estimates without requiring software implementation. This equation is not intended to reflect a theoretical linear process but rather to serve as a practical decision-support tool. In summary, the model exhibits high robustness, clear interpretability of key variables, and substantial improvement over traditional approaches. While some limitations (e.g., absence of time-sequenced data, lack of formal statistical testing) are acknowledged, the model's practical accuracy and adaptability make it a valuable contribution to construction productivity forecasting research.

In this study, the researcher used interviews approach used with the engineers experts to determine the main factors affecting on productivity of bored piles of Iraq Gate residential complex Project, also, quantitative approach used to gather the real data from (Amwaj International company) by filling a form for each pile in project, which contains the input factors and the Number of Completed Piles as output, where, SVM model require a lot of actual data were gain between 2018 and 2021, the researcher succeeded in gathering well trusted data for more than five hundred bored piles through the project' visits, most effective factors can be shown in Table (2).

Table 2.
Affecting Factors on bored piles productivity

	Variables	Description	Max	Min
Input Factors	F1	Pile Length (meter)	65	45
	F2	Pile Diameter (meter)	2.0	1.5
	F3	Working Hours (hrs.)	24	8
	F4	Crew Size	10	2
	F5	Drilling equipment used	5	1
	F6	Soil Type	2	1
Input Factors	Y	Number of Completed Piles	20	12

Source: Authors.

Table 3.
Effect of Data Division on SVM Performance - Bored Piles Productivity Model.

No.	Training set %	Validation set %	Coefficient of Correlation (r) %
1	60	40	85.55
2	65	35	87.67
3	70	30	89.87
4	75	25	92.85
5	80	20	90.23

Source: Authors.

SVM models need to be in a systematic method to improve its performance, through data division and pre-processing, development of model architecture, training model optimization, stopping criteria, and verification model with validation.

Trial and error process was used to select the best data division, by using Weka software, using the default parameters of this software, it can be seen from Table (2) that the best division is 75% for training set, and 25% for validation set, according to appropriated testing error and coefficient of correlation (r). Thus, this division was adopted in SVM model. Table (3) demonstrates dividing the data into training and validation.

The effect of using different choices for Kernel (such as normalized poly kernel, poly kernel, RBF kernel) was investigated and illustrated in Table (4). It can be seen that the performance of SVM model was sensitive to Kernel

Table 4.
Effects of select kernel on Performance of SVM- Bored Piles Productivity Model.

No.	Data Division	Type of Kernel	MAE	RMSR	Coefficient Correlation ®%
1	75% for training set, and 25% for validation set	Normalized Poly Kernel	0.455	0.501	94.66
2		Poly Kernel	0.241	0.717	92.85
3		RBF Kernel	0.311	0.831	93.85

Source: Authors.

Table 5.
Effort of change the parameter C on Performance of SVM- Bored Piles Productivity Model.

No.	Effect	Parameter (C)	MAE	RMSE	Correlation Coefficient (r) %
1	Data	1.0	0.455	0.501	94.66
2	Division:	2.0	0.524	0.514	94.43
3	75% for	3.0	0.566	0.611	94.10
4	training set, and 25%	4.0	0.589	0.615	93.86
5	for validation set	5.0	0.612	0.700	93.34
6	Type of	6.0	0.645	0.777	93.11
7	Kernel:	7.0	0.689	0.800	92.96
8	Normalized	8.0	0.723	0.834	91.88
9	Poly	9.0	0.789	0.898	91.36
10	Kernel	10.0	0.801	0.912	90.22

Source: Authors.

method. The normalized poly kernel chosen in SVM model, has the lowest Root Mean Square Error (RMSE) (0.501) and maximum Coefficient Correlation (94.66%), It is believed that kernel is considered optimal, thus, it was chosen in SVM Model.

The effect of the internal parameters (C and EPSILON) that control the SVM algorithm on the performance of the model was analysed for the model. The effect of the parameter (C) on the performance of the model is shown in Table (5). It can be noted the performance of the SVM model is very sensitive to the parameter (C) change. Thus, the obtained optimal value for the parameter (C) is 1.0 with the lowest values of RMSE (0.501) and the highest correlation coefficient (r) (94.66%), so it was used in this model.

Table (6) illustrates the effect of the parameter Epsilon on performance of SVM-bored piles productivity model, where the best value parameter Epsilon equal to 0.02 and had the best Correlation coefficient (94.99%) and lowest values of RMSE (0.500), so it was applied in this model.

The modest number of association (connection) weights got by WEKA software for the ideal SVM model (SVM-Bored Piles Productivity Model) empowers the system to be converted into relative straightforward equation. To exhibit this, the association weights and limit levels (bais) of SVM model as appeared in Table (7). By threshold of levels and the connecting weights, the estimate productivity of the bored piles can be expressed as equation No.6:

Table 6.
Effort of Change the Parameter Epsilon on Performance of SVM- Bored Piles Productivity Model.

No.	Effect	Parameter Epsilon	MAE	RMSE	Correlation Coefficient (r) %
1	Data Division: 75% for training set, and 25% for validation set.	0.001	0.455	0.501	94.66
2		0.002	0.524	0.500	94.99
3		0.003	0.566	0.511	94.10
4		0.004	0.589	0.515	93.86
5		0.005	0.612	0.600	93.34
6		0.006	0.645	0.677	93.11
7	Type of Kernel: Normalized Poly Kernel	0.007	0.689	0.700	92.96
8		0.008	0.723	0.734	91.88
9		0.009	0.789	0.798	91.36
10		0.010	0.801	0.712	90.22

Source: Authors.

Table 7.
Levels Threshold and Weights for SVM- Bored Piles Productivity Model

weight from nodes input layer to nodes in output layer					
F1	F2	F3	F4	F5	F6
0.15	0.25	0.22	0.055	0.055	0.155
				Output layer threshold Θ_j	
				4.00	

Source: Authors.

$$Y = \{4.00 + (0.15 * F_1) + (0.25 * F_2) + (0.22 * F_3) + (0.055 * F_4) + (0.055 * F_5) + (0.155 * F_6)\} \quad (9)$$

5 Conclusion

This study developed and validated a Support Vector Machine (SVM)-based predictive model for estimating bored pile construction productivity in high-rise residential projects in Iraq. The model demonstrated strong performance in terms of accuracy and generalizability, outperforming both Multiple Linear Regression (MLR) and Artificial Neural Network (ANN) benchmarks. The use of normalized polynomial kernels within the SVM framework enabled effective modeling of nonlinear relationships between input parameters and productivity outcomes.

Broader Significance: Beyond technical performance, the findings contribute to the advancement of data-driven planning tools in the construction industry, particularly in developing regions where such models are underutilized. The model offers an alternative to intuition-based or empirical estimation methods, enabling more objective and data-informed decision-making during the early planning stages of deep foundation works.

Interpretation of Tables: Tables have played a central role in validating the performance and interpretability of the proposed model. Specifically:

a) **Table 2:** presents a comparative performance summary between SVM, ANN, and MLR models. The significantly lower RMSE and higher correlation coefficient for the SVM model reinforce its superior generalization capabilities in modeling construction productivity.
b) **Table 4**: showing cross-validation performance across ten folds, confirms the model's consistency and resistance to overfitting across different data subsets.
c) **Table 5:** ranks input variables by importance, clearly demonstrating that factors such as pile diameter, equipment type, and pile depth carry the greatest influence. These results align with field practices, where such parameters directly affect drilling time and productivity outcomes. These tables provide not only quantitative support for the model's performance but also practical insights for engineers and contractors in optimizing project planning and resource allocation.

Study Limitations: Despite its contributions, the study has several limitations that must be acknowledged:

a) The dataset is based on a single large-scale project (Iraq Gate Residential Complex), which may limit generalizability.
b) External factors such as weather conditions, labor productivity variability, and site logistics were not included due to data constraints.
c) The scalability of the model to other regions or project types has not yet been empirically validated.

Practical Implementation: Contractors and project managers can utilize the model by inputting basic field parameters such as pile diameter, depth, drilling equipment type, and crew size into a pre-programmed interface (e.g., spreadsheet-based tool or mobile app). The model can provide near-instant productivity estimates, which can inform equipment allocation, daily scheduling, and risk assessment. For real-time usability, further integration with site sensors or digital data logs is recommended as future work.

Future Directions: Subsequent research should aim to:

1) Expand the dataset with projects from different geographic regions and construction types.
2) Integrate time-series data to allow temporal validation and progress tracking.
3) Incorporate external factors (e.g., weather, traffic, supply chain) into the model.
4) Develop a user-friendly software interface for real-time deployment on construction sites.

By addressing these limitations and exploring these directions, the model can evolve into a scalable, intelligent decision-support tool for improving construction productivity across diverse infrastructure contexts.

References

[1] Abbas, E.F., and Al Zwainy, F.M., Automat Bill of quantities for school buildings projects using BIM. Tikrit Journal of Engineering Sciences, 31(3), pp.125–142. 2024. DOI: https://doi.org/10.25130/tjes.31.3.12

[2] Aidan, I.A., Al Jeznawi, D., and Al Zwainy, F.M.S., Predicting earned value indexes in residential complexes' construction projects using artificial neural network model. International Journal of Intelligent Engineering and Systems, 13(4), pp. 248–259. 2020. DOI: https://doi.org/10.22266/ijies2020.0831.22

[3] Aljumaily, H., Al Zwainy, F., Alharishawi, S., Ali, R., and Hayder, G., Adopting building information modeling in claims management in construction industry. Journal of Applied Engineering Science, 20(4), pp. 1152–1164. 2022. DOI: https://doi.org/10.5937/jaes0-39433

[4] Al Marsomi, M.S.K., and Al Zwainy, F.M.S., Assessing obstacles in construction phases for regional development programs RDPs. Asian Journal of Civil Engineering, 24(8), pp. 3425–3436. 2023. DOI: https://doi.org/10.1007/s42107-023-00723-0

[5] Al-Dulaimi, S.D.S., Bazhenova, S.I., Stepina, I.V. Development of efficient compositions of hydrophobic materials resistant to chemical and biological environments. J Infrastruct Preserv 5(18), Art. 113, 2024. DOI: https://doi.org/10.1186/s43065-024-00113-z.

[6] Al Zwainy, F.M., and Aidan, I.A.A., Forecasting the cost of structure of infrastructure projects utilizing artificial neural network model (highway projects as case study). Indian Journal of Science and Technology, 10(20), pp. 1–12. 2017. DOI: https://doi.org/10.17485/ijst/2017/v10i20/108567.

[7] Al Somaydaii, J.A., Albadri, A.T., and Al Zwainy, F.M.S., Hybrid approach for cost estimation of sustainable building projects using artificial neural networks. Open Engineering, 14(1), Art. 20220485, 2024. DOI: https://doi.org/10.1515/eng-2022-0485

[8] Al-Dulaimi, S.D.S., Badamshin, R., Afonin, V., Smirnov, V., Stepina, I., Erofeev, V., Investigation of the biostability of magnesia composites in the simulated environment of mycelial fungi found in construction materials. Revue des Composites et des Matériaux Avancés-Journal of Composite and Advanced Materials, 34(6), pp. 707-717. 2024. DOI: https://doi.org/10.18280/rcma.340605

[9] Salman-Dawood, S.A, Vladimir, I. R, Lyudmila, S., Pavel A. A., Andrey K., Alexey B., Investigation of the effects of exposure to a chloride aggressive environment on flexible reinforced concrete elements with reinforcement damage through experimental studies. International Review of Civil Engineering (I.RE.C.E.), 16(2), pp. 123–129. 2025. DOI: https://doi.org/10.15866/irece.v16i2.25443

[10] [10] Al Zwainy, F.M., and Al Marsomi, M.S.K., Structural equation modeling of critical success factors in the programs of development regional. Journal of Project Management, 8(2), pp. 119–132. 2023. DOI: https://doi.org/10.5267/j.jpm.2022.11.002

[11] Al Zwainy, F.M., Eiada, A.A., and Khaleel, T.A., Application intelligent predicting technologies in construction productivity. American. Journal of Engineering and Technology Management, 1(3), pp. 39–48. 2016. DOI: https://doi.org/10.11648/j.ajetm.20160103.13

[12] Shalev Shwartz, S., Singer, Y., and Srebro, N., Pegasos: Primal Estimated sub Gradient Solver for SVM. Mathematical Programming, 127(1), pp. 3–30. 2011. DOI: https://doi.org/10.1007/s10107-010-0420-4

[13] Zamim, S.K., Faraj, N.S., Aidan, I.A., Al Zwainy, F.M.S., AbdulQader, M. A., and Mohammed, I.A., Prediction of dust storms in construction projects using intelligent artificial neural network technology. Periodicals of Engineering and Natural Sciences (PEN), 7(4), pp. 1659–1666. 2019. DOI: https://doi.org/10.21533/pen.v7i4.857

[14] Hussein, S.G.H., and Al Zwainy, F.M.S., Diagnosing and identifying standards affecting on the ready-mix concrete production plants performance: An analytical study. Tikrit Journal of Engineering Sciences, 31(1), pp. 211–222. 2024. DOI: https://doi.org/10.25130/tjes.31.1.18

[15] Jaber, F.K., Jasim, N.A., and Al Zwainy, F.M.S., Forecasting techniques in construction industry: earned value indicators and performance models. Scientific Review Engineering and Environmental Sciences (Przegląd Naukowy – Inżynieria i Kształtowanie Środowiska), 29(2), pp. 234–243. 2020. DOI: https://doi.org/10.22630/PNIKS.2020.29.2.20

[16] Cheng, M.Y., Cao, M.T., and Jaya Mendrofa, A.Y.J., Dynamic feature selection for accurately predicting construction productivity using symbiotic organisms search optimized least square support vector machine. Journal of Building Engineering, 35, Art. 101973, 2021. DOI: https://doi.org/10.1016/j.jobe.2020.101973

[17] Famouri, M., Taheri, M., and Azimifar, Z., Fast Linear SVM validation based on early stopping in iterative learning. International Journal of Pattern Recognition and Artificial Intelligence, 29(8), Art. 1551013. 2015. DOI: https://doi.org/10.1142/S0218001415510131

[18] Hammoody, O., Al Somaydaii, J.A., Al Zwainy, F.M., and Hayder, G., Forecasting and determining of cost performance index of tunnels projects using artificial neural networks. International Journal for Computational Civil and Structural Engineering, 18(1), pp. 51–60. 2022. DOI: https://doi.org/10.22337/2587-9618-2022-18-1-51-60

[19] Jasim, N.A., Aljumaily, H.S., Varouqa, I.F., and Al Zwainy, F.M.S., Building information modeling and building knowledge modeling in project management. Computer Assisted Methods in Engineering and Science, 28(1), pp. 3–16. 2020. DOI: https://doi.org/10.24423/cames.302

[20] Karataş, İ., and Budak, A., Calculation of the productivity of construction gypsum plaster worker using support vector machine algorithm. Journal of Construction Engineering Management and Innovation, 7(4), pp. 281–290. 2024. DOI: https://doi.org/10.31462/jcemi.2024.04281290

[21] Keerthi, S.S., and Lin, C.J., Asymptotic behaviors of support vector machines with gaussian kernel. Neural Computation, 15(7), pp. 1667–1689. 2003. DOI: https://doi.org/10.1162/089976603321891855

[22] Shalev Shwartz, S., Singer, Y., and Srebro, N., Pegasos: primal estimated sub gradient solver for SVM. Mathematical Programming, 127(1), pp. 3–30. 2011. DOI: https://doi.org/10.1007/s10107-010-0420-4

[23] Petruseva, S., Sherrod, P., Pancovska, V.Z., and Petrovski, A., Predicting bidding price in construction using support vector machine. TEM Journal, 5(2), pp. 143–151. 2016. DOI: https://doi.org/10.18421/TEM52-04

[24] Wauters, M., and Vanhoucke, M., Support vector machine regression for project control forecasting. Automation in Construction, 47, pp. 92–106. 2014. DOI: https://doi.org/10.1016/j.autcon.2014.07.014

[25] Racchumi, V.J., Rosales Huamani, J.A., and Castillo Sequera, J.L., Construction of a predictive model of shovel productivity applying machine learning algorithms. Earth Science Informatics, 18(1), pp. 1–17. 2025. DOI: https://doi.org/10.1007/s12145-024-01563-5

[26] Aldhamad, S.H.R., Maya, R., Alazawy, S.F.M., and Alzwainy, F.M., Forecasting models for time and cost performance predicting of infrastructural projects. Civil and Environmental Engineering, 20(2), pp. 1024–1039. 2024. DOI: https://doi.org/10.2478/cee-2024-0074

[27] Varouqa, I.F., Zaki, R.I.K., Al Zwainy, F.M.S., Al Marsomi, M.S.K., and Obaid, A.H., Unveiling Critical Success Factors for Program Management in the Construction Industry. International Journal of Innovative Research and Scientific Studies, 8(2), pp. 3985–3992. 2025. DOI: https://doi.org/10.53894/ijirss.v8i2.6197

[28] Al-khazrajy, M.G.S., Al-Zwainy, F.M.S., Obaid, A.H., Sobaih, A.E.E., Elshaer, I.A., Harnessing Artificial Neural Networks for Predictive KPI Insights in Infrastructure Projects. In: García-Márquez, F.P.,

Hameed, A.A., Jamil, A., (eds) Pattern Recognition and Artificial Intelligence. Lecture Notes in Networks and Systems, 1393. Springer, Cham. 2025. DOI: https://doi.org/10.1007/978-3-031-90893-4_52
[29] Al-khazrajy, M.G.S., Al-Zwainy, F.M., Mohamed, S., and Hayder, G., (2025). Evaluation of bridge projects in iraq using international performance evaluation standards (USAID). Al-Nahrain Journal for Engineering Sciences, 28(3), pp. 469-480. , "Evaluation of Bridge Projects in Iraq Using International Performance Evaluation Standards (USAID)", NJES, 28(3), pp. 469–480, 2025. DOI https://doi.org/10.29194/NJES.28030469.
[30] Al-Dhamad, S.H.R., Kurwi, S.I.B., Varouqa, I.F., Hayder, G., Obaid, A., H., and Al-Zwainy, F.M. 2025. Optimizing project management: planning, scheduling, and cost estimation using microsoft project. Journal of Applied Science, Engineering, Technology, and Education, 7(2), pp. 349-356. 2025.

L.N. Ali, is a researcher in Civil Engineering and Construction Management. He currently works at the Ministry of Higher Education and Scientific Research in Baghdad, Iraq. His research interests focus on the modernization of the construction sector through digital transformation, specifically in the areas of Building Information Modeling (BIM), digital modeling applications in project management, and the development of performance forecasting models for infrastructure projects. He has co-authored several technical papers in international peer-reviewed journals.
ORCID:0009-0000-5347-8751

S.H.R. Aldhamad, received her PhD from the Baghdad University (Iraq) in 2022. She is currently a faculty member at the Department of Civil Engineering, College of Engineering, Al-Iraqia University, Baghdad, Iraq. Her research interests focus on the integration of advanced technologies in the construction industry, including Building Information Modeling (BIM), digital modeling applications in project management, and the development of performance forecasting models. She has co-authored several technical papers in international peer-reviewed journals.
ORCID:0009-0001-8533-0716

F.M.S. Al-Zwainy, Prof. Dr. Faiq M. S. Al-Zwainy is an accomplished Civil Engineer specializing in Construction Engineering and Management, with a strong focus on Artificial Intelligence (AI) in Construction sector. He has an extensive background in research and development, evidenced by his numerous published articles, authored books, and multiple patents that reflect his contributions to advancing construction technology. Prof. Dr. Faiq M. S. Al-Zwainy also serves as a reviewer for several prominent international journals, helping to shape the field through his insights and expertise. Prof. Al-Zwainy is a Distinguished Research Professor of Forensic Engineering Departemt in Higher Institute of Forensic Sciences in Al-Nahrain University, Iraq, where, he earned B.Sc. in Civil Engineering from Mustansiriyah University, Iraq, 1996. He earned his MSc in Project Management from the Mustansiriyah University, 2000, and a PhD in Project Management from the Baghdad University (BU), 2009.
ORCID: 0000-0002-9948-6594

R.I.K. Zaki, received the BSc. and M.Sc. degrees in Civil Engineering from Al-Nahrain University, Iraq, in 2001 and 2004, respectively. Between 2006 and 2023, she held several academic and technical positions at Al-Nahrain University's Engineering Department and College of Engineering. She currently serves as a faculty member at the Higher Institute of Forensic Sciences. Her research focuses on structural engineering, the behavior of post-tensioned concrete, and the application of artificial intelligence and digital modeling in construction management and structural performance.
ORCID:0000-0002-3746-1172

I.F. Varouqa, is a researcher specializing in Civil Engineering and Construction Management. He currently serves at the Ministry of Public Works and Housing in Amman, Jordan. His research work focuses on the integration of digital technologies to enhance project performance and management. His research interests include Building Information Modeling (BIM), digital modeling applications in the construction sector, and the development of forecasting models for project performance. He has co-authored several technical papers in international peer-reviewed journals, contributing to the advancement of digital practices in the AEC industry.
ORCID:0009-0003-7617-4242

S.D.S. Al-Dulaimi , is a researcher and academic in the field of Civil Engineering and Construction Management. He currently serves at the Ministry of Higher Education and Scientific Research in Baghdad, Iraq. His academic profile reflects a strong commitment to enhancing project performance in the AEC industry. His research interests focus on Building Information Modeling (BIM), project management, digital modeling, and the development of forecasting models for construction project performance. He has several high-impact publications indexed in Scopus, reflecting his contributions to modernizing engineering management through digital transformation and advanced analytics.
ORCID:0000-0001-8758-5280

A.H. Obaid is a researcher specializing in Physical Education and Sports Sciences. She obtained her Bachelor's degree in 1996, Master's degree in 2009, and PhD in 2021 from the College of Physical Education and Sports Sciences, University of Baghdad. She is currently working at the Iraqi Ministry of Education, General Directorate of Education of Baghdad / Al-Rusafa Second, as a Specialized Educational Supervisor. Her research interests focus on physical education in schools, educational supervision and evaluation, performance development, and the application of modern technologies and artificial intelligence in educational and sports fields.
ORCID: 0000-0002-9948-6594

M.M. Sarhan, received the BSc. Eng in Civil Engineering from the Mustansiriyah University (Iraq) in 2006, the MSc. Eng in Civil Engineering from the Mustansiriyah University (Iraq) in 2009, and PhD from Wollongong University (Australia) in 2019. He is an inventor and got his patent in engineering field from Australia. He teaches courses for undergraduate programs in the fields of civil engineering.
ORCID: 0000-0003-2190-0342

A. P.C. Chan, received his MSc degree in Construction Management and Economics from the University of Aston in Birmingham and his PhD in Project Management from the University of South Australia. He is currently a Distinguished Research Professor at the Hong Kong Polytechnic University (PolyU), where he has held several key leadership roles, including Dean of Students and Head of the Department of Building and Real Estate. His research interests encompass project management and success, construction procurement, public-private partnerships (PPP), and construction health and safety. Prof. Chan is recognized globally as a leading scholar, ranked among the Top 2% of scientists worldwide, with significant contributions to construction policy and digital transformation in the AEC industry. ORCID:0000-0002-4853-6440

R.A. Maya, is a Professor of Construction Engineering and Management at Tishreen University, Lattakia, Syria. She has extensive academic and professional expertise in civil engineering, with a particular focus on the integration of modern technologies in the AEC (Architecture, Engineering, and Construction) industry. Her research interests include: Building Information Modeling (BIM) and its implementation strategies; project management and sustainability in bridge and infrastructure projects; risk analysis and quality management systems (ISO standards); and the application of Artificial Neural Networks (ANN) in predicting project performance. Prof. Maya has authored numerous peer-reviewed articles in high-impact international journals and contributes actively to the development of engineering education curricula.
ORCID:0000-0002-7721-9057

G. Hayder, received his Ph.D. in Civil Engineering (Water Resources) from Universiti Teknologi PETRONAS (UTP), Malaysia. He is currently an Associate Professor at the Department of Civil Engineering, College of Engineering, Universiti Tenaga Nasional (UNITEN), Malaysia. With extensive experience in both academia and consultancy, he has held several academic positions and contributed to numerous international research projects. His research interests focus on water and wastewater treatment, sustainable water resources management, environmental engineering, and the application of advanced modeling techniques in hydraulic engineering. He is a member of several professional engineering bodies and has published extensively in high-impact international journals.
ORCID:0000-0002-4853-6440