Predicting Abrasive Wear in Excavator Bucket Teeth using Cutting-Edge Vector Machines and Artificial Neural Networks*

[1]G. Ativor, [1]V. A. Temeng and [1]Y. Y. Ziggah
[1]University of Mines and Technology, P. O. Box 237, Tarkwa

Ativor, G., Temeng, V. A. and Ziggah, Y. Y. (2024), "Predicting Abrasive Wear in Excavator Bucket Teeth using Cutting-Edge Vector Machines and Artificial Neural Networks", *Ghana Mining Journal*, Vol. 24, No. 1, pp. 92-106.

Abstract

The mining industry faces a growing need to accurately predict abrasive wear of excavator bucket teeth in order to establish effective maintenance policies. Hence, developing robust predictive models that can effectively track the deterioration of ground cutting tools in harsh operational environments is a logical strategy to address this challenge. This research compared the effectiveness of three vector machine models for predicting abrasive wear of excavator bucket teeth: the least squares support vector machine (LS-SVM), the relevant vector machine (RVM), and the support vector machine (SVM). The research included a comparison of the LS-SVM model's predictive performance to that of three reference ANN techniques: radial basis function network, backpropagation neural network, and generalised regression neural network. To determine the most effective technique, the prediction results of these methods were assessed using metrics such as mean square error (MSE), root-mean-square error (RMSE), correlation coefficient (R), coefficient of determination (R^2), mean absolute error (MAE), and Nash-Sutcliffe Efficiency (NSE). Additionally, the Bayesian Information Criterion (BIC) was employed to select the best performing predictive method. According to the evaluation results, the LS-SVM model outperformed both the RVM and SVM methods and the reference ANN techniques, as it achieved the lowest values for MSE, RMSE, and MAE (0.025726, 0.160394, and 0.131220) respectively, along with the highest values for R, R^2, and NSE (0.999900, 0.999800 and 0.999794) respectively. Moreover, LS-SVM attained the lowest BIC value (-300.5774), demonstrating its superior ability to predict on-site wear of excavator bucket teeth.

Keywords: Artificial Neural Networks, Excavator Bucket teeth, Abrasive Wear, Least Squares Support Vector Machine

1 Introduction

In surface mining operations, large excavators are used to excavate and load materials into transportation units. Excavator buckets are fitted with cutting components because of the varying characteristic composition of the material being excavated. During operation, the bucket teeth, which are the cutting elements, are the first to come in contact with the formation and, as a result, exposed to dynamic-resistive forces and intensive abrasive wear. As the excavating bucket teeth moves relatively against the muck, the interacting surfaces are progressively removed, leading to abrasive wear (Bayer, 2004). Exposure to such conditions over an extended period results in the gradual deterioration of the bucket teeth, which can eventually cause deformations or breakages in severe cases. The loading and energy consumption of an excavator are both adversely affected when the bucket teeth are in such a state, because they are unable to achieve their primary function. Consequently, the efficiency of the excavator decreases, and production becomes less economically viable due to the increased costs associated with replacement of the worn-out cutting elements and reduced output. To optimise the bucket teeth for excavation purposes, it is essential to have a comprehensive understanding of the excavation wear properties of the material being excavated (Miletić *et al*., 2018; Plinninger, 2010), the machine or tool characteristics (Kumar and Alam, 2016) and the field operational parameters (Khorzoughi and Hall, 2016a; Prakash *et al*., 2013).

Excavator bucket teeth used in harsh working conditions are typically made of alloy steel and include additional durable compounds such as chromium, molybdenum, and nickel to extend the tool life against abrasive wear (Suryo *et al.*, 2018; Fernandez *et al.*, 2001). Prior studies have attempted to carry out experiments in the laboratory in order to fabricate and investigate alternative materials for the production of excavator bucket teeth, to monitor the heat treatment on new products, and to estimate the filler materials used for hard facing the cutting elements exposed to abrasive wear (Keleş and Yildirim, 2020; Suryo *et al*., 2018). However, the challenge with employing laboratory abrasive wear studies is that these tests do not entirely and precisely mimic field conditions (Blickensderfer, 1988). Attempts have also been made to assess the interaction effects on ground penetrating tools and to analyse wear during operation by using numerical approaches like the Discrete Element Method (DEM) and the Finite Element Method (FEM) (Choudhry, 2020; Curry and Deng, 2017; Kumar and Alam, 2016). Nonetheless, the computational procedures involved in these methods are very costly (Arregui-Mena *et al*., 2016; Rackl and Hanley, 2017). Researchers such as Bhushan (2000), She *et al.* (2022), Bošnjak *et al.* (2018), Khorzoughi and Hall (2016b), and Dong *et al.* (2023) have attributed the inefficiency of current approaches to

*Manuscript received February 13, 2024
Revised version accepted June 24, 2024
https://doi.org/10.4314/gm.v22i1.3

their failure to consider real parameters like the quartz content of the muck being excavated, uniaxial compressive strength, excavator loading cycles, and blasted rock sizes, all of which contribute to the wear of excavator bucket teeth. Addressing the issue of monitoring excavator bucket teeth wear is a significant concern for industry players. Currently, no mechanism or methodology in literature has adopted these field-controllable factors to forecast the on-site abrasive wear of excavator bucket teeth. Therefore, this study focuses on utilising these site-controllable parameters to predict the abrasive wear of excavator bucket teeth.

The mining industry requires reliable and accurate prediction outcomes for making decisions in their maintenance practices. This need has become more pronounced, leading to an increased demand for the testing of new techniques that can provide such results. Artificial intelligence techniques have emerged as a cutting-edge and potent computational approach for developing intelligent systems in recent years (Zhang and Lu, 2021). Due to the exceptional tolerance for randomness, as well as the robust learning capabilities, artificial intelligence (AI) techniques are gaining attention but have less been explored to predict abrasive wear of excavator bucket teeth. The lone AI technique applied in the literature that utilised vibration signals to forecast the wear of excavator bucket teeth in surface excavation processes is the adaptive neuro-fuzzy inference system (Miletić *et al*., 2020). Hence, the authors conducting this research propose that it would be intellectually enriching to explore the potential of AI tools for predicting the abrasive wear of excavator bucket teeth.

Selecting suitable artificial intelligence techniques is essential for accurate predictions in diverse mining environments. For the first time, this research evaluates three vector machine models: Least Squares Support Vector Machine (LS-SVM), Relevance Vector Machine (RVM), and Support Vector Machine (SVM), each offering unique benefits, for the prediction of excavator bucket teeth wear. LS-SVM is valued for its computational efficiency (Wang and Hu, 2005; Temeng *et al*., 2022), RVM for its sparse solutions and enhanced interpretability (Zio and Di Maio, 2012), and SVM for its robustness across various applications (Roy and Chakraborty, 2023). Assessing these models in a Ghanaian surface mine, considering essential excavation wear factors such as uniaxial compressive strength, quartz content in rock fragments, rock fragment sizes resulting from blasting, and the duration of the excavator's loading activities, is particularly important to ensure the results are relevant and applicable to similar mining settings. Subsequently, the accuracy of the most suitable vector machine model is compared with that of widely used ANN predictive models, namely the Radial Basis Function Neural Network, the Backpropagation Neural Network, and the Generalised Regression Neural Network.

2 Resources and Methods Used

2.1 Resources

The study area where the investigation took place is an open pit mine that can be found in Ghana. Specifically, the site is located about 133 km west of Koforidua, the regional capital and approximately 3 km west of New Abirem, the district capital. It is also about 180 km northwest of Accra (Kaba, 2013). The mine sits on a deposit at the extreme northern and southeastern borders of the Greenstone Belt. In the northwestern part of the mining area, Tarkwaian sediments, which consist of conglomerate, sandstone, and phyllite, unconformably overlay the Birimian volcanic belts. Fig. 1 depicts the location of the study site.

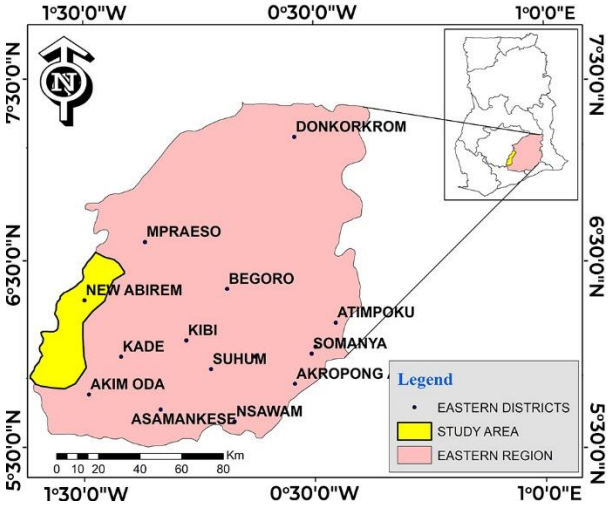

Fig. 1 Study Area

The mining operations involve a fleet of equipment, which comprises of two Liebherr 9400 hydraulic shovels and two hydraulic backhoes for muck handling. Additionally, there are eighteen 785 Caterpillar rear dump trucks available at the mine, each capable of carrying up to 134 tonnes of ore to the stockpile or crusher, or transporting non-economical materials to the waste dump. The loading process employs a single backup loading system and removes a 9 m bench using 3 m sequential flitch excavations. The mine has seven Pantera drill rigs, out of which five have a 165 mm drill bit diameter, and they are primarily utilised for drilling blast holes. The remaining two drill rigs have 115 mm diameter of bit designs, and are mainly used for drilling wall control holes. The mining site also employs a staggered drill pattern with a 4.0 m × 4.0 m burden and spacing in the ore zone and a 5.2 m × 5.2 m burden and spacing in the waste zone.

2.2 Data Description and Pre-processing

With the objective of creating predictive models, a total of 579 historic data points over a span of 289.5 effective working days, were acquired from the mine. The data comprised of variables, such as bucket scooping duration (mins), equivalent quartz content (%), mean fragment sizes (cm), and uniaxial compressive strength (MPa) as the input parameters and average wear of excavator bucket teeth (mm) as the output parameter. In this study, the FRAGTrack^{TM} system, a cutting-edge technology made up of smart image processing device was fixed on top of the excavator rollover protection structure (ROPS). This system was used to record and report the size distribution of rock fragments in both two-dimensional and three-dimensional formats. The mean fragment sizes gathered from each recording were utilised for the model development. The Liebherr 9400 excavator considered for this study, was operated in a zone which comprised of phyllite, sandstone and greywacke rock types. Compression tests were performed to obtain uniaxial compressive strength (UCS), following the protocols outlined by the International Society for Rock Mechanics (Aydin, 2015). The equivalent quartz content (EQC) was acquired by performing X-ray diffraction laboratory test. Global Positioning System (GPS), typically referred to as dispatch system, was used to monitor the excavator's cycle time, with a specific focus on extracting datapoints related to the duration of bucket scooping activity. Ground Engaging Tool (GET) measurement and change-out tracker was used to monitor the longitudinal dimensions of bucket teeth as they gradually diminished towards the bucket lip at the end of each shift. The average wear of the five teeth positions was used for analysis in this study. Table 1 presents the statistical information pertaining to the dataset that was gathered. The correlation coefficient matrix described in Table 2 shows the variability between the dependent and response variables..

Table 1 Statistical Interpretation of Data

Parameters	Unit	Min	Max	Average	Std Dev
Digging Duration	min	164.28	407.60	298.24	57.86
Fragment Sizes	cm	12.16	41.82	22.20	6.34
Uniaxial Compressive Strength	MPa	58.00	157.00	87.06	40.75
Equivalent Quartz Content	%	24.00	60.00	35.40	15.73
Average Wear	mm	24.00	56.00	36.08	10.85

Table 2 Correlation Coefficient Matrix of the Dataset

Parameters	Loading Duration	Mean Fragment Sizes	UCS	Equivalent Quartz Content	Average Wear
Loading Duration	1				
Mean Fragment Sizes	-0.5923	1			
Uniaxial Compressive Strength	-0.9006	0.6121	1		
Equivalent quartz content	-0.8737	0.6042	0.9952	1	
Average wear	-0.8079	0.4908	0.9688	0.9603	1

The data was separated into two sets: the testing dataset and the learning (training) dataset. The learning dataset comprised 80% of the entire dataset which was employed to construct the predictive models while the second set (20%) was reserved for testing the efficacy of the candidate techniques. To accomplish this, the hold-out cross-validation strategy was adopted in the data division criterion. This criterion entails using a training data size that is greater than the size of testing data (Kohavi, 1995). Also, the learning and testing data points should be carefully sorted out in order to represent the population's characteristic features (Dobbin and Simon, 2011). Not complying with this condition may result in overfitting, particularly when the hold-out sample falls outside the domain of the training dataset. To ensure uniformity in the dataset and reduce the impact of values with large magnitudes on smaller ones, the input parameters in the dataset used for model development were normalised using Equation (1) (Ali *et al*., 2014). This was necessary because the parameters had varying physical units and dimensions. Normalising the parameters to the range of [-1, 1] helped to achieve quick and better convergence to a global optimal result.

$$n_i = \frac{\left(x_i - x_{\min}\right)\times(y_{\max} - y_{\min})}{(x_{\max} - x_{\min})} + y_{\min} \qquad (1)$$

n_i is data point after normalisation, where the observed data point is represented by x_i, and x_{min} and x_{max} define the minimum and maximum values of the observed points. In this context, y_{min} and y_{max} are marked values of -1 and 1, respectively.

2.3 Methodology

In the first part of this section, the mathematical descriptions, and procedures for constructing the vector machine models explored and applied in this research are presented. The second section of the research methodology provides a concise conceptual overview of three commonly utilised neural networks namely, Radial Basis Function Neural Network (RBFNN) Generalised Regression Neural Network (GRNN), and Backpropagation Neural Network (BPNN). The procedure for building each model is also provided.

2.3.1 Support Vector Machine

This study employed the SVM to perform function approximation by using input vector $x = (x_1, x_2, x_3, x_4, ..., x_R)$ and correlating target value (*f*) (Cortes and Vapnik, 1995). To accomplish this, an SVM classifier was applied to regression $f(x)$, as shown in Equation (2).

$$f(x) = b + (\phi(x) \cdot w) \tag{2}$$

weight vector $w = (w_1, w_2, w_3, w_4, ..., w_R)$, the bias term *b*, and the basis function vector $\phi(x) = (\phi_1(x_1), ..., \phi(x_k))$ referring to a sequence of non-linear transformations, are used in the SVM. The *R*-dimensional input vector is transformed from a low-dimensional state to a high state using the dynamic basis function $\phi(x)$ .The support vector machine seeks to determine the optimal separating hyperplane that maximises learning capability of dataset to solve a specific problem (Brezak *et al*., 2012).
To solve Equation (2), Equation (3) is required, subject to the restrictions specified in Equation (4). It is worth noting that the SVM's network is not predetermined, and the input variables that support the network or structure are generally chosen during the model learning process (Malakar *et al*., 2018).

$$\underset{w,b,\xi,\xi^*}{\text{minimise}} \frac{1}{2} \| w \|^2 + C \sum_{J=1}^{R} (\xi_j + \xi_j^*) \tag{3}$$

Subject to:

$$\left\{ \begin{array}{c} y_j - \left(b + w^{\mathrm{T}} \phi(x_j)\right) \le \xi_j + \varepsilon \\ b - y_j + w^{\mathrm{T}} \phi(x_j) \le \xi_j^* + \varepsilon \\ \xi_j^*, \xi_j \ge 0 \end{array} \right\} j = 1, 2, 3, 4, ..., R \tag{4}$$

The positive trade-off parameters, ε and C control the scale of the empirical error in maximisation, while the slack variables ξ and $\xi*$ solve the learning errors based on the loss expression in Equation (4).

Typically, the quadratic expression described by Equation (3) with respect to Equation (4) can be computed by duality approach, as shown in Equation (6) with respect to Equation (7), by maximising the multipliers and Lagrangian function in Equation (5).

$$\begin{aligned} L = {} & \frac{1}{2} \| w \|^2 + C \left(\sum_{J=1}^{r} \xi_j^* + \sum_{J=1}^{r} \xi_j \right) \\ & - \sum_{J=1}^{r} \alpha_j \left(w^{\mathrm{T}} \phi(x_j) - (y_j - \xi_j) + (b + \varepsilon) \right) \\ & - \sum_{J=1}^{r} \alpha_j^* \left(y_j - b - w^{\mathrm{T}} \phi(x_j) \le \xi_j^* + \varepsilon \right) \\ & - \sum_{J=1}^{r} \left(\gamma_j \xi_j + \gamma_j^* \xi_j^* \right) \end{aligned} \tag{5}$$

The Lagrange multipliers $\gamma_j^*, \gamma_j, \alpha_j^*, \alpha_j$ are used in conjunction with the Lagrangian, *L*. To solve Equation (5), it is necessary to comply with Karush-Kuhn-Tucker requirements. Here, entries that satisfy the requirements by having Lagrangian multipliers that are not zeros are considered to be support vectors since they aid the classifier or estimator's structure (Wu, 2007).

$$\begin{aligned} \text{maximise } & -\frac{1}{2} \sum_{k,j}^{r} (\alpha_k - \alpha_k^*)(\alpha_j - \alpha_j^*) N_{kj} \\ & + \sum_{j}^{r} y_j (\alpha_j - \alpha_j^*) - \varepsilon \sum_{j}^{r} (\alpha_j - \alpha_j^*), \end{aligned} \tag{6}$$

$$\text{with respect to } \left\{ \begin{array}{c} \sum_{j}^{r} \left(\alpha_j - \alpha_j^* \right) = 0 \\ \alpha_j^*, \alpha_j \in [0, C] \end{array} \right\} \tag{7}$$

$N_{jk} = K(x_j, x_k) = \phi(x_j)^{\mathrm{T}} \phi(x_k)$ is the kernel technique (Sánchez, 2003). The positivity restrictions of $\gamma_j^*, \gamma_j, \alpha_j^*, \alpha_j \ge 0$ must be achieved by the dual variables in Equation (6). Following this, the values of α_j and α_j^* obtained

from the solution given by Equation (6) subject to Equation (7) are utilised in the model hyperplane to calculate the optimal weight vector, *w*. expressed in Equation (8).

$$w = \sum_{j}^{r} \left(\phi(x_j)\right)(\alpha_j - \alpha_j^*) \quad (8)$$

As a result, Equation (9) can then be employed to express the SVM classifier or estimator on the model regression.

$$f(x) = \sum_{j}^{r} (\alpha_j - \alpha_j^*) K(x, x_j) + b \quad (9)$$

$K(x, x_j)$ used in the Equation (9) is described by (Fadel *et al*., 2016) as the kernel function. Here, the polynomial kernel function was utilised because it has the potency to effectively recognise the dynamic pattern in the training data. Equation (10) provides the mathematical definition for the polynomial kernel function.

$$K(x_j, x_k) = \left(1 + x_j^T x_k\right)^q \quad (10)$$

2.3.2 Least Squares Support Vector Machine

To provide an alternative to the classical support vector machine formulation, (Suykens and Vandewalle, 1999) introduced the LS-SVM. The LS-SVM model is constructed based on a given learning dataset, *M*, $\{x_i, y_i\}_{i=1}^{M}$, an input dataset, $x_i \in R^M$, and an output value, $y_i \in r$. Here, R^M is a vector space with *M*-dimensions and r is a single vector space. The LS-SVM model is formulated with Equation (11) in feature space as:

$$y(x) = w^{\mathrm{T}} \phi(x) + b \quad (11)$$

To overcome the maximisation challenges presented in Equation (12), a non-linear approach of mapping function is employed to map the model parameters to a feature space with higher dimension, along with an adjustable weight vector (*w*), a transpose operation (*T*), and a scalar cutoff point (*b*). Also, to attain the objective, Wang and Hu (2005) indicated that the equality conditions described in Equation (13) should be satisfied.

$$\text{minimise } \tfrac{1}{2} w^{\mathrm{T}} w + \gamma \tfrac{1}{2} \sum_{i=1}^{M} e_i^2 \quad (12)$$

$$y(x) = b + e_i + w^T \phi(x_i), \quad i = 1, 2, 3, ..., M \quad (13)$$

the error variable e_i and the regularisation variable γ, which control the balance between minimising the fitting function and achieving flatness, are both important in this context. Equation (14) which is the Lagrangian function *L* is employed to compute Equation (12) while adhering to the requirements provided in Equation (13).

$$L(w, b, \alpha, e) = \tfrac{1}{2} w^{\mathrm{T}} w + \gamma \tfrac{1}{2} \sum_{i=1}^{M} e_i^2 - \sum_{i=1}^{M} \alpha_i \left[(b + w^{\mathrm{T}} \phi(x)) - y_i + e_m \right] \quad (14)$$

The α_i values represent the Lagrange multipliers, while the partial derivative of *L* subject to the terms, presented in Equation (15) was computed to meet the optimality criterion stated in Equation (13).

$$\begin{gathered} w = \sum_{i=1}^{M} \alpha_i \phi(x_i) \Rightarrow \frac{\partial L}{\partial w} = 0 \\ \sum_{i=1}^{M} \alpha_i = 0 \Rightarrow \frac{\partial L}{\partial d} = 0 \\ \frac{\partial L}{\partial e_i} = 0 \Rightarrow \gamma e_i = \alpha_i, \quad \forall i = 1, 2, 3, ..., M \\ \frac{\partial L}{\partial a_i} = 0 \Rightarrow y_i = b + w^{\mathrm{T}} \phi(x_i) + e_i, \ \forall i = 1, 2, 3, ..., M \end{gathered} \quad (15)$$

By removing the *e* and *w* variables, the linear Karush-Kuhn-Tucker requirement (Equation (16)) is obtained.

$$\begin{bmatrix} 1_M^p & 0 \\ \gamma^{-1} I_M & 1_M^p \Omega \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ Y \end{bmatrix} \quad (16)$$

where $1_M = [1, ..., 1]^p$, $Y = [y_1, ..., y_M]^p$ and $\alpha = [\alpha_1, ..., \alpha_N]^p$ while 1_M is an $M \times M$ identity matrix with the kernel parameter (Ω) described in Equation (17).

$$v_{ij} = \phi(x_i)^p \left(\phi(x_j)\right) = T(x_i, x_j) \quad (17)$$

The kernel function used was the linear kernel function in this research, Equation (18).

$$\mathrm{T}(x_i, x_j) = x_i(x_j^{\mathrm{p}}) \quad (18)$$

The output of α and *b* can be determined using Equation (15). As a result, the LS-SVM's predictive outcomes are given in Equation (19).

$$y(x) \sum_{I=1}^{M} \alpha_i \, \mathrm{T}(x_i, x_j) + b \quad (19)$$

In order to run the LS-SVM effectively, the parameter γ in Equation (12) must be tuned to yield the optimum result and this was achieved through iterative process.

2.3.3 Relevance Vector Machine

The RVM is described as a probabilistic sparse kernel technique that uses a defined distribution for parameter weights which are influenced by a

combination of hyperparameters (Tipping, 2001). The theory underlying RVM is based on linear regression principles (Zio and Di Maio, 2012). As expressed in Equation (20), the RVM regression model is formulated to be:

$$y_k = \varepsilon_k + f(x_k) \tag{20}$$

the noise factor in the dataset is represented by ε_k. When using a linear model, function $f(x)$ is converted into a distribution based on a defined function $\varphi_k(x)$, as shown in Equation (21).

$$f(x) = \sum_{k=1}^{M} w_k \cdot \varphi_k(x) \tag{21}$$

where $w = (w_1, w_2, w_3, ..., w_M)$ represents the weight vector used for detecting boundaries. The Equation (21) can be modified into vector form (Equation (22)).

$$y = \Phi w + \varepsilon \tag{22}$$

where σ^2 represents the variance in the dataset, the design matrix, Φ, has dimension $M \times (M + 1)$, and is formed by the T^{th} row vector as defined in Equation (23). ε is the noise component with a mean, zero.

$$\Phi_k(x_m) = \left[1, T(x_m, x_1), T(x_m, x_2), ..., T(x_m, x_M)\right] \tag{23}$$

By utilising the formulation given in (Equation (24)), it becomes possible to represent the probability of the learning data set as:

$$p(y \mid w, \sigma^2) = ((\sigma^2 . 2\pi)^{M-2}) \exp\left\{ \| y - \Phi w \| . - \frac{1}{2\sigma^2} \right\} \tag{24}$$

During the training stage of the relevance support machine, weight parameter (w) is bounded by applying a Gaussian defined function with zero mean on it, as shown in Equation (25) (Kong *et al*., 2019).

$$p(w \mid \alpha) = \prod_{t=1}^{N} M(w_t \mid 0, \alpha_t^{-1}) \tag{25}$$

where $a_k = (a_1, a_2, a_3, ..., a_N)$, is employed to describe the reciprocal variance of weight, w_t. Bayes's principle can be utilised to express the posterior distribution of all uncertain variables as shown in Equation (26):

$$p(w, \alpha, \sigma^2 \mid y) = p(y \mid w, \alpha, \sigma^2) p(w, \alpha, \sigma^2) \frac{1}{p(y)} \tag{26}$$

here, $p(y)$ can be described in Equation (27) as:

$$p(y) = \iiint p(y \mid w, \alpha, \sigma^2) p(w, \alpha, \sigma^2) dw \, d\alpha \, d\sigma^2 \tag{27}$$

Since the normalisation of the integral in Equation (27) cannot be directly computed, $p(w, \alpha, \sigma^2)$, which represents the posterior in Equation (26), cannot be solved immediately. Instead, Equation (28) is employed to solve the $p(w, \alpha, \sigma^2)$.

$$p(w, \alpha, \sigma^2 \mid y) = p(w \mid, \alpha, \sigma^2, y) p(\alpha, \sigma^2 \mid y) \tag{28}$$

The application of Bayes' rule results in the expression of the posterior function over the parameter weights, which is represented in Equation (29).

$$p(w \mid \alpha, \sigma^2 \mid y) = \frac{p(w \mid \alpha) \cdot p(y \mid w, \sigma^2)}{p(y \mid \alpha, \sigma^2)} \sim M(n, \Sigma) \tag{29}$$

In Equation (30) and (31), n and Σ represent the average and covariance, respectively.

$$n = \sigma^{-2} \Sigma \Phi^{\mathrm{T}} y \tag{30}$$

$$\Sigma = (\mathrm{A} + \sigma^{-2} \Phi^{\mathrm{T}} \Phi)^{-1} \tag{31}$$

where $\mathrm{A} = \mathrm{diag}(\alpha) = \mathrm{diag}\,(\alpha_0, \alpha_1, \alpha_2, ..., \alpha_M)$.

Equation (32) expresses the probabilistic model over the learning targets, which is derived by computing the parameter weights to determine the probability of the hyperparameters (Yu *et al*., 2004).

$$p(y / w, \alpha^2) = \int \left(p(y / w, \sigma^2) \right) p(w / \alpha) dw \sim M(0, \mathrm{F}) \tag{32}$$

In Equation (33), C represents the covariance matrix.

$$\mathbf{F} = \sigma^{-2} \mathbf{I} + \Phi \mathbf{A}^{-1} \Phi^{\mathrm{T}} \tag{33}$$

The logarithmic probability function of the learning targets is represented by Equation (34):

$$\begin{aligned} \ln \mathrm{p}(\mathrm{y}/\alpha, \sigma^2) = & \frac{M}{2} In(\sigma^{-2}) - \frac{1}{2}(\sigma^{-2} y^T y - n^T \Sigma^{-1} n) \\ & - \frac{M}{2} In(2\pi) + \frac{1}{2} \sum_{k=0}^{M} In(\alpha_k) \end{aligned} \tag{34}$$

hence, the variable weights, w, is provided by the average of probability function in Equation (29), and the hyper-parameters can be obtained by maximising the value attained from Equation (30). Given a different input, x_{new}, the probability

function of the prediction value, y_{new}, can be derived using Equation (35), which is written as:

$$p(y_{new} \mid x_{new}, \hat{\alpha}, \hat{\sigma}^2) = \int p(y_{new} \mid x_{new}, w, \hat{\sigma}^2)... \\ p(w \mid y, \hat{\alpha}, \hat{\sigma}^2)dw \sim M(n_{new}, \sigma^2_{new}) \quad (35)$$

where n_{new} in Equation (36) and σ^2_{new} in Equation (37) represent the average and corresponding variance of the predictive variables.

$$n_{new} = n^{\mathrm{T}}\Phi(x_{new}) \quad (36)$$

$$\sigma^2_{new} = \hat{\sigma}^2 + \Phi(x_{new})^{\mathrm{T}}\Sigma\Phi(x_{new}) \quad (37)$$

2.3.4 Procedure for Developing SVM, LS-SVM and RVM Models

The abrasive wear prediction model for the SVM was constructed using the first order of polynomial kernel. The hyper-parameters, C as well as ε, were tuned using the conventional iterative trial-and-error approach to enhance the model's performance (Tseng *et al*., 2016).

Equation (12) used the simplex search methodology to fine-tune the regularisation parameter (γ) for the LS-SVM (Ziggah *et al*., 2022) to achieve optimal values.

The optimum value for the Gaussian kernel width variable or kernel bandwidth (σ) was iteratively obtained for the relevance vector machine model, and this also plays a crucial role in determining the model's performance in this current research.

2.3.5 Backpropagation Neural Network

BPNN is a feed forward neural network with three connected layers. BPNN, with the three connected layers comprising of an input layer or unit, layer of hidden neurons, and output unit was utilised. The model architecture for a back propagation neural network is shown in Fig. 2, wherein the input unit receives external inputs or information and transfers them into the network for processing. After receiving the input(s), each neuron within a hidden unit undergoes a computational non-linear activation process to transform the inputs (Xu *et al*., 2015). The two commonly applied transfer functions for this purpose are the hyperbolic tangent and the logarithmic function. The response or output unit, at the extreme end of the architecture, features a linear neuron and receives the processed information from the hidden unit, which is then transmitted to an external receptor (Dorofki *et al*., 2012).

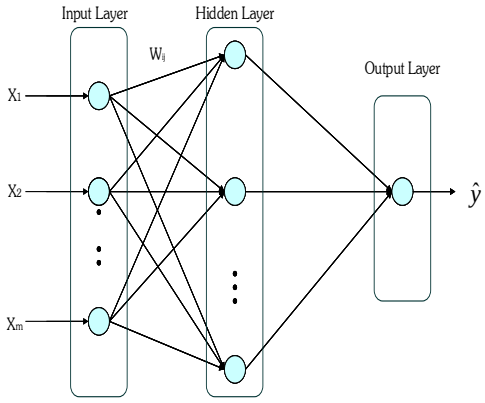

Fig. 2 BPNN Architecture

The momentum coefficient used during the training of a neural network can significantly impact its convergence speed and stability. A high momentum coefficient can expedite convergence but could also result in instability if set too high. Conversely, a low momentum coefficient can lead to sluggish training and difficulties avoiding local minima (Fukuoka *et al*., 1998). For the wear prediction of excavator bucket teeth, a 0.03 learning rate and 0.8 momentum coefficient were employed during the 5000-epoch training process in this study. Insights into the mathematical concept of BPNN development have been provided in literature (Yegnanarayana, 2009). It has been demonstrated that one layer of hidden neurons is capable for a BPNN to function as a universal approximator of intricate problems (Ziggah *et al*., 2016). Hence, this research utilised one layer of hidden neurons. The feedback from the hidden unit, denoted as Equation (38), is subsequently transferred to the output unit.

$$Y_i = g\left(\sum_{j=1}^{m} w_{ij}X_j + b_i\right) \quad (38)$$

where g ( ) represents the transfer function implemented by the hidden unit, b_i denotes the bias term, and w_{ij} is the weight values. Using Equation (39), the predictor-response operation is performed using linear activation equation in the output unit to generate the final results, $\hat{y}$.

$$\hat{y} = Y_i \quad (39)$$

2.3.6 Radial Basis Neural Network

RBFNN has three units as well: an input unit, a unit with hidden neurons, and an output unit. The structure contains various components such as the vector transformation of input data $A_i = (A_1, A_2, A_3, A_4, ..., A_m)^T$, radial basis functions $(\phi_1, \phi_2, \phi_3, \phi_4, \ldots, \phi_L)$, parameter weights as $w_i = (w_1, w_2, \ldots, w_L)$, and $\hat{y}$ as the final output. In this network, the input layer sends external data to the hidden units, without any weight computations. The hidden unit is comprised of neurons that use a radial basis as activation function. Here, the function is responsible for quasi-linear processing activity within the unit (Shin and Park, 2000).

Gaussian radial basis function has proved to be the most efficient approximation used in RBFNN (Fasshauer and McCourt, 2012), therefore it was applied in this study. Each neuron figures out the Euclidean interval between the input objects and the focal point of the Gaussian distribution. To successfully apply the RBFNN, it is crucial to identify appropriate focal points or centres for the Gaussian function. The Gaussian distribution function is defined by two different parameters: the centre (c_j) and the width parameter (σ_j). The output net_j is given in Equation (40).

$$net_j = \exp\left(-\frac{\| X_i - c_j \|^2}{2\sigma_j^2}\right) \quad (40)$$

where the calculated Euclidean distance between c_j and X_i is represented by $\| X_i - c_j \|$. The hidden neuron outputs are first of all weighted and summed to create inputs for the output or response layer. The output unit applies a linear transformation function to the input, leading to the final result $\hat{y}_p$ of the RBFNN as shown in Equation (41).

$$\hat{y}_p = a + \sum_{j=1}^{m} w_{jp} net_j \quad (41)$$

where *a* represents the bias term, w_{jp} is the weight utilised to describe the connection between the hidden unit and the final unit, and *m* refers to the count of neurons employed in the hidden unit. Fig. 3 illustrates a RBFNN architecture.

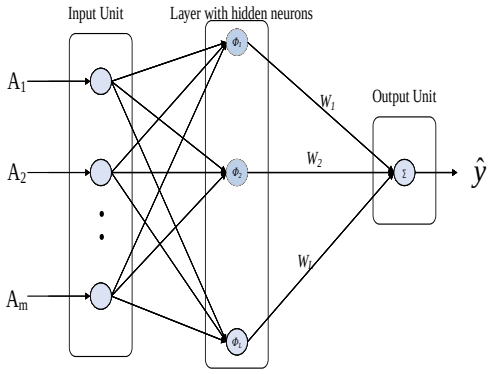

Fig. 3 RBFNN Architecture

During the RBFNN learning process, the parameters that are adjusted include the centres, width parameters, and weights. It should be emphasized that the training process is focused on minimising the errors between the observed data a_p and the predicted results y_p, (Equation (42)).

$$\text{Minimise (MSE)} = \frac{1}{k}\sum_{p=1}^{k}(a_p - \hat{y}_p)^2 \quad (42)$$

here, *k* refers to the number of observed points and MSE refers to mean square error.

2.3.6 Generalised Regression Neural Network

This neural network employs a single pass learning technique and features four different units: pattern unit, input unit, output unit, and the summation unit (Specht, 1991). This structure is shown in Fig. 4. The input unit takes external data and forwards it to the pattern unit. The pattern neurons utilise a Gaussian function for transferring information subject to a spreading index (Dong-xiao *et al.*, 2008). The D-summation neurons within the summation unit calculate the unweighted outcomes, whereas S-summation neurons compute the weighted outcomes of the pattern neurons. The predicted results $Y(x)$ are obtained by multiplying the S-summation outputs by the inverse of D-summation outputs (Tasdemir *et al.*, 2013). This is presented in Equation (43).

$$Y(x) = \sum_{n=1}^{m} w_k \exp\left(-\frac{(x - x_n)^2}{2\psi^2}\right) \times \left(\sum_{k=1}^{m} \exp\left(-\frac{(x - x_n)^2}{2\psi^2}\right)\right)^{-1} \quad (43)$$

here, the input vector m represents the element count, ψ refers to the parameter used for spreading, and x and x_k define the *kth* element of the vector.

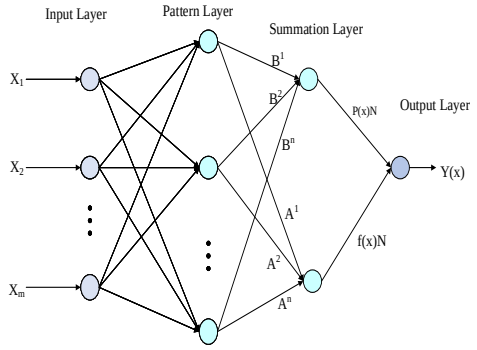

Fig. 4 GRNN Architecture

2.3.7. Procedure for Developing BPNN, RBFNN and GRNN

The development of backpropagation neural network model involved selecting the learning algorithm, activation functions, and number of hidden neurons and units. Here, a BPNN with an input unit, an output unit, and one hidden unit were used. Through a series of iterations, the suitable count of neurons for the hidden unit was achieved, and the function for activation implemented in the hidden unit was the hyperbolic tangent, while in the output unit a linear function was employed for activation processes (Karlik and Olgac, 2011). The BPNN training was achieved by implementing the Levenberg-Marquardt trainer.

The performance of RBFNN depends on the optimum count of neurons found in the hidden unit, and the width parameter for developing the model. Hence, the optimal RBFNN model was constructed by determining the optimum count of hidden neurons and the width parameter iteratively.

The GRNN's performance is regulated by the width parameter, which was also achieved through iterative fine-tuning.

2.4 Performance Metrics

In this study, artificial intelligence models, namely LS-SVM, RVM, SVM, BPNN, RBFNN, and GRNN models, were evaluated and compared for their suitability. To determine the effectiveness of the various techniques constructed, the research used six statistical indicators, including mean square error (MSE), root-mean-square error (RMSE), correlation coefficient (R), coefficient of determination (R^2), mean absolute error (MAE), and Nash-Sutcliffe Efficiency (NSE). The mathematical expressions for these indicators are presented in Equation (44) through (49).

$$\mathrm{MSE} = \frac{1}{m} \times \sum_{i=1}^{m} (a_i - p_i)^2 \tag{44}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{m} \times \sum_{i=1}^{m} (a_i - p_i)^2} \tag{45}$$

$$\mathrm{R} = \frac{\sum_{i=1}^{m} (a_i - a_{av})(p_i - p_{av})}{\sqrt{\sum_{i=1}^{m} (a_i - a_{av})^2} \times \sqrt{\sum_{i=1}^{m} (p_i - p_{av})^2}} \tag{46}$$

$$\mathrm{R}^2 = \left(\frac{\sum_{i=1}^{m} (a_i - a_{av})(p_i - p_{av})}{\sqrt{\sum_{i=1}^{m} (a_i - a_{av})^2} \times \sqrt{\sum_{i=1}^{m} (p_i - p_{av})^2}} \right)^2 \tag{47}$$

$$\mathrm{MAE} = \frac{1}{m} \times \sum_{i=1}^{m} | a_i - p_i | \tag{48}$$

$$\mathrm{NSE} = 1 - \left(\frac{\sum_{i=1}^{m} (a_i - p_i)^2}{\sum_{i=1}^{m} (a_i - a_{av})^2} \right) \tag{49}$$

where a_{av} is the mean of the observations, p_{av} is the mean of the predicted outcomes, and m represents the sum of observations. p_i and a_i are the predicted wear values and observed wear values, respectively. In order for a model to be considered a good predictor, it is expected to produce relatively low values for MSE and RMSE. When the values of RMSE and MSE are low, it can be inferred that the model has a low residual error. R, R^2 and NSE are used to assess the linear relationships between predicted and observed bucket teeth wear values, with a value close to one indicating a highly significant relationship.

2.5 Model Selection

The Bayesian Information Criterion (BIC) was employed to identify the best wear prediction model among the candidate predictors (Schwarz, 1978).

The BIC factors in the size of the test data, response values, number of model parameters, and actual values. The candidate model or technique with the lowest BIC score is considered the most suitable for accurate wear prediction. Equation (50) provides the mathematical formulation of the BIC.

$$\mathrm{BIC} = m\ \mathrm{In}\left(\frac{1}{m}\times\sum_{i=1}^{m}(A_i - R_i)^2\right) + \gamma \mathrm{In}(m) \qquad (50)$$

where R_i denotes the response values, m denotes the number of model samples, A_i defines the observed values, and γ denotes the number of model features or parameters.

3 Results and Discussion

3.1 Developed Vector Machine Models

The study involved developing three different vector machine learning models: SVM, LS-SVM, and RVM for bucket teeth wear prediction. The development of the SVM method based on the values of ε and C yielding the performance of the model, and this was achieved with optimal values of 0.000000001 and 65, respectively. The LS-SVM model was also developed based on γ value, which was obtained to be 63 998.7101. Finally, the RVM model was developed by adjusting the kernel bandwidth, resulting in an optimal value of 1.0744. Table 3 presents the results of testing evaluation metrics for each of the three vector machine models used in the study.

Table 3 Results of Testing Evaluation Metrics for the Vector Machine Models

Models	R	R^2	MAE	RMSE	MSE	NSE
LS-SVM	0.999900	0.999800	0.131220	0.160394	0.025726	0.999794
RVM	0.999576	0.999152	1.058132	1.161269	1.348546	0.989218
SVM	0.999884	0.999768	0.135159	0.172130	0.029629	0.999763

3.1.1 Comparison of LS-SVM, RVM and SVM models

The error metrics determined the extent to which the response values of the models diverge from the actual observations. Although the results (Table 3) imply that there are no significant differences among the three vector machine methods, it is essential to choose a model with high accuracy and precision for a robust analysis of bucket teeth wear. In terms of accuracy, the LS-SVM technique exhibited the lowest MAE, RMSE, and MSE values of 0.131220, 0.160394 and 0.160394, respectively. Again, LS-SVM exhibited superior performance on the testing dataset, achieving the highest values of R (0.999900), R^2 (0.999800), and NSE (0.999794) compared to RVM and SVM. The results from Table 3 suggest that the LS-SVM has excellent ability to capture patterns effectively from the learning data samples and adapt effectively to the testing dataset. Consequently, it was chosen as the most appropriate technique for predicting wear of hydraulic excavator bucket teeth. Additionally, the LS-SVM technique's strength can be attributed to its inherent flexibility to implement the linear Karush–Kuhn–Tucker requirement, which provides a robust computational capability for handling non-linearity challenges in datasets (Zhang and Zhang, 2016).

3.2 Developed Artificial Neural Networks

The BPNN model was achieved by using a combination of hyperbolic tangent in the hidden unit and a linear function for activation processes in the output unit or layer, along with implementing the Levenberg-Marquardt trainer. The optimal BPNN structure was [4-24-1], which means that it had four input variables in the input unit, twenty-four hidden neurons, and one response in the output unit.

For the RBFNN model, the study findings indicate that the width parameter of 0.6 attained the best results. The structure of the optimal RBFNN model was [4-25-1], which also indicates four inputs, 25 hidden neurons, and one response.

Through a series of iterations, the optimal value for the width parameter was identified, and in this study, a value of 0.10 was found to be the most effective for the GRNN model since it resulted in the least amount of errors and also avoided overfitting.

3.2.1 Comparison of LS-SVM model and Reference Neural Networks

This section involves comparing the LS-SVM vector machine model with reference ANN techniques such as BPNN, RBFNN, and GRNN to determine how well it predicts the wear of excavator bucket teeth, which has been studied in previous sections.

Table 4 summarises the statistical test results that evaluated the performance of each method using the testing dataset.

Table 4 Evaluation Metrics for Predictive Models on Test Dataset

Model	Evaluation Metrics					
	MSE	RMSE	MAE	R	R^2	NSE
LS-SVM	0.025726	0.160394	0.131220	0.999900	0.999800	0.999794
BPNN	0.166901	0.408535	0.334227	0.999376	0.998657	0.998666
RBFNN	0.217623	0.466500	0.327321	0.999144	0.998289	0.998260
GRNN	0.064539	0.254046	0.207838	0.999746	0.999492	0.999484

In this comparative analysis, the error indicators MSE, RMSE and MAE, were utilised to measure the predictability of the models and to determine the extent of deviation between the response values and the observed values. The model accuracy was considered to be excellent if the MSE, RMSE and MAE values approached zero. Here, LS-SVM technique exhibited the lowest values of MSE, RMSE and MAE, representing 0.025726, 0.160394 and 0.131220, respectively, indicating its superior predictive performance compared to the other models. The R value, representing the percentage of prediction accuracy, was also utilised to evaluate the effectiveness of the models. The LS-SVM model outperformed all other models with the highest R value (0.999900), followed by GRNN with an R value (0.999746). The remaining models had R values below 0.9997, as shown in Table 4. The R^2 and NSE values serve as useful metrics for evaluating the agreement between predicted and observed data, with R^2 ranging from zero to one and NSE ranging from negative infinity to one. Higher values for these metrics indicate a better match between predicted wear values and observed wear values. The results of the R^2 and NSE analysis presented in Table 4 suggest that the LS-SVM method outperformed the other methods. This is evident from the remarkably high R^2 and NSE values of 0.999800 and 0.999794, respectively, indicating a better fit for abrasive wear prediction. Thus, these findings strongly support the assessment that the LS-SVM method is the most appropriate AI technique for predicting bucket teeth wear compared to the ANN methods. The results presented in Table 4 are visually displayed in Figs. 5 through 10, which demonstrate that LS-SVM is superior to the applied ANN models.

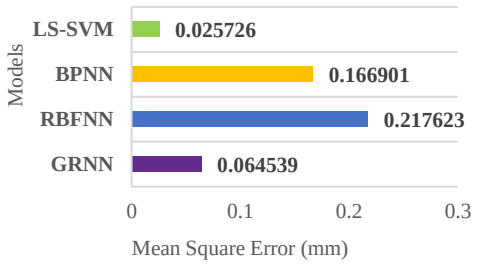

Fig. 5 Mean Square Error Results for the Applied Models

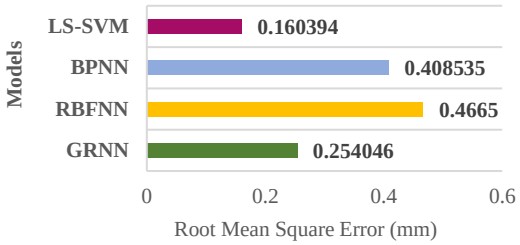

Fig. 6 Root Mean Square Error Results for the Applied Models

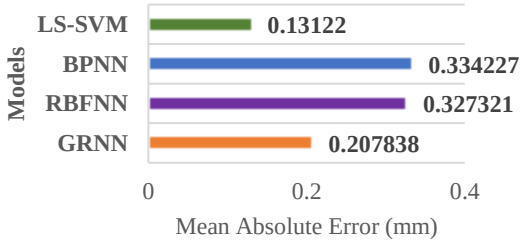

Fig. 7 Mean Absolute Error Results for the Applied Models

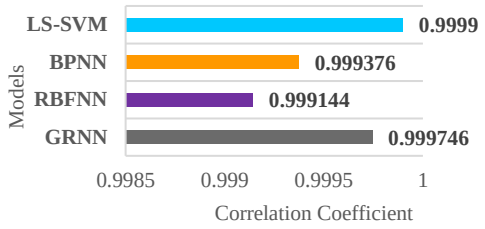

Fig. 8 Correlation Coefficient Results for the Applied Models

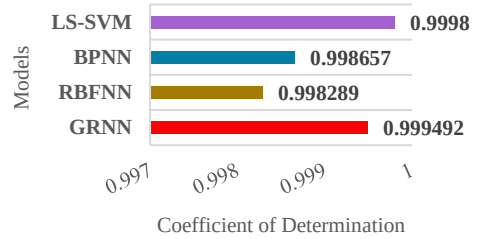

Fig. 9 Coefficient of Determination Results for the Applied Models

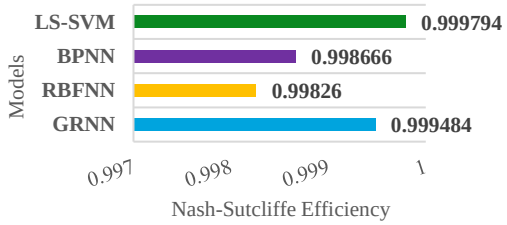

Fig. 10 Nash-Sutcliffe Efficiency Results for the Applied Models

3.3 Selection of Best Performing Model

To select the most effective model from the investigated methods, the Bayesian Information Criterion (BIC) was adopted. The model or technique with the lowest BIC value is deemed the best candidate. According to the BIC results shown in Table 5, it is evident that the LS-SVM approach outperforms the other methods in predicting the abrasive wear of excavator bucket teeth in the study area. The LS-SVM model, having the lowest BIC value of -300.5774 among the considered methods (Table 5), supports this assessment, making it the preferred technique over RVM, SVM, BPNN, RBFNN, and GRNN methods. This comparison is further visualized in Fig. 11.

Table 5 Calculated BIC Values for the Applied Methods

Models	BIC Value
LS-SVM	-300.5774
SVM	-257.6058
RVM	-162.5752
BPNN	-137.8974
RBFNN	-114.8107
GRNN	-220.5581

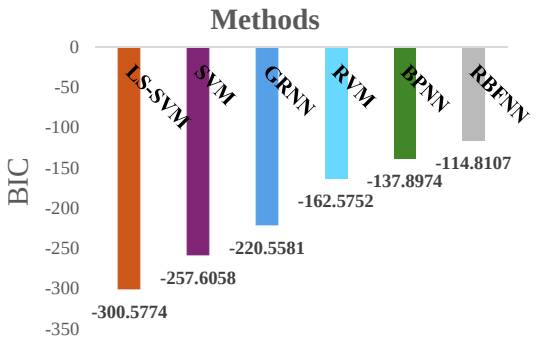

Fig. 11 BIC Results of the Applied Methods

4 Conclusions

This research examined the effectiveness of three different vector machine schemes (LS-SVM, RVM, and SVM) for predicting excavator bucket teeth wear. The effectiveness of the techniques was assessed using six different statistical metrics including MAE, MSE, RMSE, R, R^2, and NSE. The results of the statistical measures showed that the LS-SVM technique was superior to both the RVM and SVM methods in terms of robustness and prediction accuracy. To determine the predictability of the LS-SVM technique, the study employed other reference methods, including BPNN, RBFNN, and GRNN, to the same dataset to ascertain their performance comparatively. The LS-SVM's superior performance is demonstrated by its lowest values of MSE (0.025726 mm), RMSE (0.160394 mm), and MAE (0.131220 mm). In contrast, the alternative methods, including RVM, SVM, BPNN, RBFNN, and GRNN, exhibited MSE, RMSE, and MAE values ranging from 0.029629 to 1.348546, 0.254046 to 1.161269, and 0.207838 to 1.058132, respectively. Furthermore, LS-SVM achieved the highest values for R (0.999900), R^2 (0.999800), and NSE (0.999794) compared to the other techniques. The BIC results further confirmed the superior performance of the LS-SVM technique, as it had the lowest calculated BIC value of -300.5774. In comparison, the prospective methods, including RVM, SVM, BPNN, RBFNN, and GRNN, exhibited BIC values ranging from -257.6058 to -114.8107. The research findings suggest that LS-SVM could serve as a valuable tool for the mining industries, where excavating and loading remain crucial. The LS-SVM method's robust calibration efficiency and adaptation capabilities make it a potent predictive tool.

Acknowledgement

The authors extend their gratitude to the Ghana National Petroleum Corporation (GNPC) for funding this research.

References

Ali, P. J., Faraj, R. H., Koya, E., Ali, P. J. and Faraj, R. H. (2014), "Data Normalisation and Standardisation: A Technical Report", *Machine Learning Technical Reports*, Vol. 1, No. 1, pp. 1-6.

Arregui-Mena, J. D., Margetts, L. and Mummery, P. M. (2016), "Practical Application of the Stochastic Finite Element Method", *Archives of Computational Methods in Engineering*, Vol. 23, pp. 171 - 190.

Aydin, A., (2008), "ISRM Suggested Method for Determination of the Schmidt Hammer Rebound Hardness: Revised Version", *International Journal of Rock Mechanics and Mining Sciences*, pp. 1-8.

Bayer, R. G. (2004), *Mechanical Wear Fundamentals and Testing*, CRC Press, New York, 416 pp.

Bhushan, B. (2000), "Tribology of Earthmoving, Mining, and Minerals Processing", In *Modern Tribology Handbook*, 2nd edition, CRC Press, pp. 1361-1400.

Blickensderfer, R. (1988), "Design Criteria and Correction Factors for Field Wear Testing", *Wear*, Vol. 122, No. 2, pp. 165-182.

Bošnjak, S. M., Arsić, M. A., Gnjatović, N. B., Milenović, I. L. and Arsić, D. M. (2018), "Failure of the Bucket Wheel Excavator Buckets", *Engineering Failure Analysis*, Vol. 84, pp. 247-261.

Brezak, D., Majetic, D., Udiljak, T. and Kasac, J. (2012), "Tool Wear Estimation Using an Analytic Fuzzy Classifier and Support Vector Machines", *Journal of Intelligent Manufacturing*, Vol. 23, pp. 797-809.

Choudhry, J. (2020), "A Study of Wear and Load Behaviour on Bucket Teeth for Heavy-Duty Cable Shovels", *Published MSc Thesis*, Lulea University of Technology, 70 pp.

Cortes, C. and Vapnik, V. (1995), "Support-vector networks", *Machine Learning*, Vol. 20, pp. 273-297.

Curry, D. R. and Deng, Y. (2017), "Optimising Heavy Equipment for Handling Bulk Materials with Adams-EDEM Co-Simulation", *Proceedings of the 7th International Conference on Discrete Element Methods*, pp. 1219-1224.

Dobbin, K. K. and Simon, R. M. (2011), "Optimally Splitting Cases for Training and Testing High Dimensional Classifiers", *BMC Medical Genomics*, Vol. 4, No. 31, pp. 1-8.

Dong, Z., Jiang, F., Tan, Y., Wang, F., Ma, R. and Liu, J. (2023), "Review of the Modeling Methods of Bucket Tooth Wear for Construction Machinery", *Lubricants*, Vol. 11, No. 6, pp. 1-19.

Dong-xiao, N., Da, L. and Mian, X. (2008), "Electricity Price Forecasting Using Generalised Regression Neural Network Based on Principal Components Analysis", *Journal of Central South University of Technology*, Vol 15, No. 2, pp. 316–320.

Dorofki, M., Elshafie, A. H., Jaafar, O., Karim, O. A. and Mastura, S. (2012), "Comparison of Artificial Neural Network Transfer Functions Abilities to Simulate Extreme Runoff Data", *International Proceedings of Chemical, Biological and Environmental Engineering*, Singapore, Vol. 33, pp. 39 – 44.

Fadel, S., Ghoniemy, S., Abdallah, M., Sorra, H. A., Ashour, A. and Ansary, A. (2016), "Investigating the Effect of Different Kernel Functions on the Performance of SVM for Recognizing Arabic Characters", *International Journal of Advanced Computer Science and Applications*, Vol. 7, No. 1, pp. 446–450.

Fasshauer, G. E. and McCourt, M. J. (2012). "Stable Evaluation of Gaussian Radial Basis Function Interpolants", *SIAM Journal on Scientific Computing*, Vol. 34, No. 2, pp. A737-A762.

Fernandez, J. E., Vijande, R., Tucho, R., Rodríguez, J. and Martín, A. (2001), "Materials Selection to Excavator Teeth in Mining Industry", *Wear*, Vol. 250 No. 12, pp. 11-18.

Fukuoka, Y., Matsuki, H., Minamitani, H., and Ishida, A. (1998), "A Modified Back-Propagation Method to Avoid False Local Minima", *Neural Networks*, Vol. 11, No. 6, pp. 1059-1072.

Kaba, F. A. (2013), "Student Internship Technical Report on Newmont Golden Ridge Limited, Akyem", *University of Mines and Technology*, Tarkwa, pp. 1-15.

Karlik, B. and Olgac, A. V. (2011), "Performance Analysis of Various Activation Functions in Generalised MLP Architectures of Neural Networks", *International Journal of Artificial Intelligence and Expert Systems*, Vol. 1, No. 4, pp. 111-122.

Keles, A. and Yildirim, M. (2020), "Improvement of Mechanical Properties by Means of Titanium Alloying to Steel Teeth Used in the Excavator", *Engineering Science and Technology, An International Journal*, Vol. 23, No. 5, pp. 1208-1213.

Khorzoughi, B. M. and Hall, R. (2016a), "A Study of Digging Productivity of An Electric Rope Shovel for Different Operators", *Minerals*, Vol. 6, No. 2, pp. 1-6.

Khorzoughi, B. M. and Hall, R. (2016b), "Diggability Assessment in Open Pit Mines: A Review", *International Journal of Mining and Mineral Engineering*, Vol. 7, No. 3, pp. 181-209.

Kohavi, R. (1995), "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection", *International Joint Conference on Artificial Intelligence (IJCAI)*, Vol. 14, No. 2, pp. 1137-1145.

Kong, D., Chen, Y., Li, N., Duan, C., Lu, L., and Chen, D. (2019) "Relevance Vector Machine for Tool Wear Prediction", *Mechanical Systems and Signal Processing*, Vol. 127, pp. 573-594.

Kumar, B. and Alam, T. (2016), "Excavator Bucket Tooth Wear Analysis", *International Conference on Electrical, Electronics and Optimisation Techniques*, pp. 3364-3365.

Malakar, P., Mukherjee, A. and Sarkar, S. (2018), "Potential Application of Advanced Computational Techniques in Prediction of Groundwater Resource of India", *Groundwater of South Asia*, pp. 643-655.

Miletic, F., Jovančić, P. and Djenadić, S. (2018), "Behaviour Determining of Bucket Wheel Drive Depending on The Wear Impact of The Cutting Elements", *Procedia Structural Integrity*, Vol. 13, pp. 1644-1650.

Miletic, F., Jovanic, P. D., Milovanevic, M. and Ignjatovic, D. (2020), "Adaptive Neuro-Fuzzy Prediction of Operation of The Bucket Wheel Drive Based on Wear of Cutting Elements", *Advances in Engineering Software*, Vol. 146, pp. 1-6.

Plinninger, R. J. (2010), "Hardrock Abrasivity Investigation Using the Rock Abrasivity Index (RAI)", *Geologically Active*, Taylor and Francis Group, London, pp. 3445 - 3452.

Prakash, A., Murthy, V. M. S. R. and Singh, K. B. (2013), "Rock Excavation Using Surface Miners: An Overview of Some Design and Operational Aspects", *International Journal of Mining Science and Technology*, Vol. 23, No. 1, pp. 33-40.

Rackl, M. and Hanley, K. J. (2017), "A Methodical Calibration Procedure for Discrete Element Models", *Powder Technology*, Vol. 307, pp. 73-83.

Roy, A. and Chakraborty, S. (2023), "Support Vector Machine in Structural Reliability Analysis: A Review", *Reliability Engineering and System Safety*, Vol. 233, pp. 109-126.

Sánchez, A. V. D. (2003), "Advanced Support Vector Machines and Kernel Methods", *Neurocomputing*, Vol. 55, pp. 5-20.

She, L., Zhang, S. R., Wang, C., Wu, Z. Q., Yu, L. C. and Wang, L. X. (2022), "Prediction Model for Disc Cutter Wear During Hard Rock Breaking Based on Plastic Removal Abrasiveness Mechanism", *Bulletin of Engineering Geology and the Environment*, Vol. 81, No. 10, pp. 432.

Shin, M. and Park, C. (2000), "A Radial Basis Function Approach to Pattern Recognition and Its Applications", *ETRI Journal*, Vol. 22, No. 2, pp. 1-10.

Specht, D. F. (1991), "A General Regression Neural Network", *IEEE Transactions on Neural Network*, Vol. 2, No. 6, pp. 568 – 576.

Suryo, S. H., Bayuseno, A. P., Jamari, J. and Ramadhan, M. A. R. (2018), "Analysis of AISI Material Power of AISI 4140 Bucket Teeth Excavator Using Influence of Abrasive Wear", *AIP Conference Proceedings*, Vol. 1977, No. 1, pp. 1 - 10.

Suykens, J. A. and Vandewalle, J. (1999), "Least Squares Support Vector Machine Classifiers. *Neural Processing Letters*, Vol. 9, pp. 293-300.

Schwarz, G. (1978), "Estimating the Dimension of a Model", *The Annals of Statistics*, pp. 461-464.

Tasdemir, Y., Kolay, E. and Kayabali, K. (2013), "Comparison of three Artificial Neural Network Approaches for Estimating of Slake Durability Index", *Environmental Earth Sciences*, Vol. 68, pp. 23-31.

Tipping, M. E. (2001), "Sparse Bayesian Learning and the Relevance Vector Machine", *Journal of Machine Learning Research*, Vol. 1, pp. 211-244.

Temeng, V. A., Arthur, C. K. and Ziggah, Y. Y. (2022), Suitability Assessment of Different Vector Machine Regression Techniques for Blast-Induced Ground Vibration Prediction in

Ghana", *Modeling Earth Systems and Environment*, Vol. 8, No. 1, pp. 897-909.

Tseng, T. L., Aleti, K. R., Hu, Z. and Kwon, Y. (2016), "E-Quality Control: A Support Vector Machines Approach", *Journal of Computational Design and Engineering*, Vol. 3, No. 2, pp. 91-101.

Wang, H. and Hu, D. (2005), "Comparison of SVM and LS-SVM for Regression, *International Conference on Neural Networks and Brain*, Vol. 1, pp. 279-283.

Wu, H. C. (2007), "The Karush–Kuhn–Tucker Optimality Conditions in an Optimisation Problem with Interval-Valued Objective Function", *European Journal of Operational Research*, Vol. 176, No. 1, pp. 46-59.

Xu, B., Zhang, H., Wang, Z., Wang, H. and Zhang, Y. (2015)", "Model and Algorithm of Backpropagation Neural Network Based on Expanded Multichain Quantum Optimisation", *Mathematical Problems in Engineering*, Vol. 2015, pp. 1-11.

Yegnanarayana, B. (2009), *Artificial Neural Networks*, Prentice-Hall of India Private Limited, New Delhi, 476 pp.

Yu, W. M., Du, T. and Lim, K. B. (2004), "Comparison of the Support Vector Machine and Relevant Vector Machine in Regression and Classification Problems", *Control, Automation, Robotics and Vision Conference*, Vol. 2, pp. 1309-1314.

Zhang, C. and Zhang, H. (2016), "Modelling and Prediction of Tool Wear using LS-SVM in Milling operation", *International Journal of Computer Integrated Manufacturing*, Vol. 29 No. 1, pp. 76-91.

Zhang, C. and Lu, Y. (2021), "Study on Artificial Intelligence: The State of The Art and Future Prospects. *Journal of Industrial Information Integration*, Vol. 23, pp. 1-7.

Ziggah, Y. Y., Issaka, Y. and Laari, P. B. (2022), Evaluation of Different Artificial Intelligent Methods for Predicting Dam Piezometric Water Level", *Modelling Earth Systems and Environment*, Vol. 8, No. 2, pp. 2715-2731.

Ziggah, Y. Y., Youjian, H., Yu, X. and Basommi, L. P. (2016), "Capability of Artificial Neural Network for Forward Conversion of Geodetic Coordinates (ϕ, λ, h) to Cartesian Coordinates (X, Y, Z)", *Mathematical Geosciences*, Vol. 48, No. 6. pp. 687 – 721.

Zio, E. and Di Maio, F. (2012), "Fatigue Crack Growth Estimation by Relevance Vector Machine", *Expert Systems with Applications*, Vol. 39, No. 12, pp. 10681-10692.

Authors

Godwin Ativor is a Postgraduate Research Assistant at the Department of Mining Engineering of the University of Mines and Technology (UMaT), Tarkwa. He holds a BSc (Hons.) degree in Mining Engineering and is currently a Doctoral candidate in Mining Engineering at UMaT. His current research interests include modelling and simulation of mine production and equipment systems, application of artificial intelligence, discrete element modelling, and drill and blast design and optimisation.

Victor Amoako Temeng is an Associate Professor in Mining Engineering. He obtained his BSc (Hons.) and PgD in Mining Engineering degrees from University of Mines and Technology (UMaT), Tarkwa. He holds MSc degree from the University of Zambia and PhD degree from the Michigan Technological University. He is a member of the Society of Mining, Metallurgy and Exploration (MSME). His areas of specialisation include Operations Research, Materials Handling and Computer Applications.

Yao Yevenyo Ziggah is a Senior Lecturer at the Geomatic Engineering Department of the University of Mines and Technology (UMaT). He holds a BSc in Geomatic Engineering from Kwame Nkrumah University of Science and Technology, Kumasi, Ghana. He obtained his Master of Engineering degree and PhD in Geodesy and Survey Engineering from China University of Geosciences (Wuhan). His research interests include artificial intelligent application in engineering, geodetic coordinate transformation, gravity field modelling, height systems and geodetic deformation modelling. frequency control of power system and electrical cable tension control.