Advances in Nano Research, *Vol. 19, No. 4 (2025) 297-317*
https://doi.org/10.12989/anr.2025.19.4.297

Nano and molecular mechanism effects of CsV vascular endothelial functions using machine learning

Lianfeng Li[1], Yanhong Yang[2], Guodao Zhang[3,4], Anwu Huang[5], Xumei Huang[5], Shanjiang Chen[5], Deyu Peng Bin Lin[*5] and Xiaojun Ji[**5]

[1]*Network Information Center, The Maternal and Child Health Hospital of Guangxi Zhuang Autonomous Region, Nanning 530002, China*
[2]*College of Computer Science and Engineering, Tianjin University of Technology, Tianjin 300384, China*
[3]*Institute of Intelligent Media Computing, Hangzhou Dianzi University, Hangzhou 310018, China*
[4]*Shangyu Institute of Science and Engineering Co.Ltd. Hangzhou Dianzi University, Shaoxing 312300, China*
[5]*Department of Cardiology, Wenzhou Central Hospital, Wenzhou 325000, Zhejiang, China*

(*Received August 20, 2022, Revised November 28, 2024, Accepted June 19, 2025*)

Abstract. Chikusetsusaponin V's (CsV) effects on Endothelial Nitric Oxide Synthase (eNOS) and vascular endothelial functions were the focus of this investigation. Various CsV doses were introduced into in vitro cultures of Bovine Aortic Endothelial Cells (BAECs), Human Umbilical Vein Endothelial Cells (HUVECs), and animal models. To evaluate the impact of CsV on endothelial activities, vascular stiffness, eNOS mRNA and protein expression, and Nitric Oxide (NO) production were examined using quantitative PCR (qPCR), Western Blotting (WB), and B-ultrasound imaging. Protein mass spectrometry, molecular docking, bioinformatics, and network pharmacology were further applied to predict upstream transcription factors and molecular interactions regulating eNOS activity. Complementarily, advanced machine learning (ML) models including neural networks, Random Forests (RF), Support Vector Machines (SVM), and RF–Fuzzy Logic were employed to predict endothelial responses under varying CsV conditions. The neural network achieved the highest predictive accuracy (86.36%), while the RF–Fuzzy Logic model demonstrated superior precision (90.91%) and recall (83.33%). Feature importance analysis identified impedance modulus and water contact angle (WCA) as critical determinants of CsV-induced endothelial regulation. These findings provide nano- and molecular-level insights into the mechanisms by which CsV modulates endothelial function, integrating experimental assays with ML-driven predictions to inform potential therapeutic strategies for cardiovascular diseases.

Keywords: cyclic shear viscosity (CSV); endothelial cell adhesion; machine learning; neural networks; random forests-fuzzy logic; support vector machines (SVMs)

1. Introduction

The heart's interior walls, blood arteries, and lymphatic vessels are lined by ECs, a thin layer of flattened cells (Trimm and Red-Horse 2023). Among their critical functions are the regulation of blood flow dynamics, the facilitation of oxygen and nutrient transport, the assistance of immune cell trafficking, and the preservation of tissue balance. Furthermore, various mediators, including vaso-constrictors and Nitric Oxide (NO), are produced and secreted by the endothelium to orchestrate the complex regulation of vascular tone (Godo and Shimokawa 2017). The main causes of endothelial dysfunction include decreased generation of vasodilators and poor function of ECs. The importance of this dysfunction as a precursor to cardiovascular illnesses is becoming more acknowledged, leading to increased attention. Measurements of the brachial artery's Flow-Mediated Dilation (FMD) have recently emerged as the gold standard for evaluating endothelial function, but other approaches exist (Flammer *et al.* 2012). Research conducted by (Inaba *et al.* 2010) and (Green *et al.* 2011) has shown that FMD may independently predict cardiovascular events. According to research conducted by (Ras *et al.* 2013), those with no prior cardiovascular issues may reduce their risk of cardiovascular disease by 4% for every 1% rise in FMD; on the other hand, hypertension and other cardiovascular risk factors are associated with a 13% decrease in risk for every 1% rise in FMD (Inaba *et al.* 2010). Thus, avoiding or at least lessening endothelial dysfunction is of utmost importance. Exercise improves endothelial function and reduces the risk of cardiovascular illnesses; it is also easy to perform and convenient. Type 2 diabetics' endothelium function was shown to be enhanced by exercise training, namely aerobic and mixed exercise, according to a meta-analysis (Qiu *et al.* 2018). Similarly, (Chen *et al.* 2023) found that concurrent exercise may improve endothelial function and decrease arterial stiffness in individuals with type 2 diabetes. Lew *et al.* (Lew *et al.* 2022) found similar results, indicating that exercise training may enhance endothelium function in healthy post-menopausal women. Notably, endothelial function in hypertensive persons may be improved with exercise training. A comprehensive analysis of endothelial function

*Corresponding author, Ph.D., Professor,
E-mail: riadmarzok06@gmail.com
**Co-corresponding author, Ph.D., Professor,
E-mail: 186743147@qq.com

ISSN: 2287-237X (Print), 2287-2388 (Online)

in prehypertensive and hypertensive individuals found that hypertensive individuals could benefit from regular moderate-intensity AE, prehypertensive individuals could benefit from High-Intensity Interval Training (HIIT), and Resistance Exercise (RE) could be an additional option for improving endothelial function (Waclawovsky *et al.* 2021). In addition, AE improved endothelial function in obese and overweight people (Cortes *et al.* 2023). In addition, endothelial function was favorably affected by AE and RE in people in good health (Shivgulam *et al.* 2023). Exercise training enhances endothelial function even in heart failure patients with a decreased ejection fraction (Pearson and Smart 2017). However, research on the impact of various forms of exercise on endothelial function in middle-aged and elderly persons has shown contradictory findings. This population may respond differentially to various forms of exercise, including AE, RE, and HIIT, concerning FMD. At this time, it is not known which kind of exercise best enhances endothelium function in middle-aged and elderly persons. The vascular endothelium is the lining of the cardiovascular system, which is continuous and just one cell thick (Amelio *et al.* 2014, Anderson *et al.* 2018). Atherosclerotic lesions are caused by EC dysfunction, which becomes apparent in sections of arterial arteries prone to disease. The intima is penetrated by blood-borne Low-Density Lipoprotein (LDL) particles, which bind to monocytes via defective ECs. Foam cells result from monocyte differentiation into macrophages (Andriani *et al.* 2016). To create fibromuscular plaques, pro-inflammatory ECs and foam cells release growth factors and chemokines, which interact with smooth muscle cells (Aviv *et al.* 2001). Changing vascular morphology causes alterations to the blood vessel's blood flow patterns. An important hemodynamic factor in EC and vascular homeostatic homeostasis is the steady or pulsing LSS (Breitling *et al.* 2004) typical of the straight artery. Anxieties, proliferative, apoptotic, anti-monocyte adhesion, and anti-thrombotic are only a few of its many actions (Clarke *et al.* 2008, Childs *et al.* 2018). The anti-atherosclerotic responses, including the expression of Endothelial Nitric Oxide Synthase (ENOS) (Du *et al.* 2010, Costarelli *et al.* 2017), transcription factors nuclear factor-E2-related factor 2 (Nrf2) (Guerrero *et al.* 2015, Gytz *et al.* 2017), and Krüppel-like factor 2 (KLF2) (Hastie and Stuetzle 1989, Hebbel *et al.* 2020), are promoted by LSS, which also inhibits monocyte adhesion and Vascular Smooth Muscle Cell (VSMC) proliferation (Coppé *et al.* 2010, Anderson *et al.* 2018). On the other hand, the OSS becomes the LSS at the carotid bifurcation or aortic arch (Hughey and Butte 2015). Microvascular and epicardial endothelial dysfunctions are directly linked to local OSS, which makes it more likely that early coronary AS will arise and progress (Jaffe *et al.* 1973). On top of that, OSS speeds up the development of AS and makes plaque more susceptible to damage by increasing the adhesion of leukocyte-ECs and inducing the release of inflammatory cytokines in ECs (Johnson *et al.* 2007, Jia *et al.* 2019). One of the main causes of AS is the activation of ECs and an increase in the permeability of the EC monolayer, both of which are linked to persistent inflammation in OSS (Jong *et al.* 2013). ML has become a transformative tool in biological research, allowing scientists to handle the vast amounts of complex data generated by modern experimental techniques (Yazdani *et al.* 2021a). ML algorithms drive discoveries across various fields, from genomics and proteomics to cellular imaging and personalized medicine, by uncovering patterns, predicting outcomes, and streamlining analyses. One significant advantage of ML in biological research is its ability to handle high-dimensional datasets, such as those from genomics or transcriptomics, where millions of data points and hundreds of variables can make manual analysis challenging (Mohammadhassani *et al.* 2014, Naghipour *et al.* 2020). ML models manage these complexities by identifying key features, filtering noise, and providing accurate and interpretable results (Naghipour *et al.* 2020). In addition to managing data, ML excels at pattern recognition, which is essential for classifying cells, identifying disease subtypes, and differentiating between healthy and diseased tissue. This capacity makes ML invaluable for diagnostics and personalized medicine, where algorithms can guide individualized treatment by predicting responses to therapies based on patient-specific data. Furthermore, ML models, intense learning techniques, are adept at revealing complex relationships within biological systems, modelling intricate interactions between genes, proteins, and environmental factors that are often challenging to decipher with traditional approaches. ML also plays a pivotal role in enhancing imaging diagnostics. Convolutional Neural Networks (CNNs) and other ML algorithms are widely used to analyze medical images, allowing for precise detection and classification of abnormalities, such as cancerous cells. This has led to more accurate diagnostics and informed decision-making in clinical settings. Moreover, in drug discovery, ML is accelerating progress by predicting the efficacy and safety of drug candidates, reducing the time and costs associated with bringing new treatments to market. ML is revolutionizing biological research by empowering researchers to interpret complex data, identify hidden patterns, and make data-driven predictions. Its applications reshape fields like genomics, medical imaging, and personalized medicine, contributing to more precise diagnostics, targeted therapies, and expedited scientific discoveries. With the increasing complexity and volume of biological data, machine learning remains crucial for advancing the understanding of life at a molecular level and improving the effectiveness of healthcare solutions. This study aims to investigate the nano and molecular mechanistic effects of CSV on vascular endothelial function, emphasising enhancing understanding through advanced ML techniques. By analyzing endothelial cell responses to CSV, this research aims to identify complex patterns in cellular adhesion strength, a critical factor in vascular health and disease development. Novelty is introduced by comparing the performance of three distinct ML models neural networks, RF, and SVMs each focusing on accurately predicting endothelial adhesion strength under varying CSV conditions. This approach contributes to vascular biology by providing a data-driven framework to

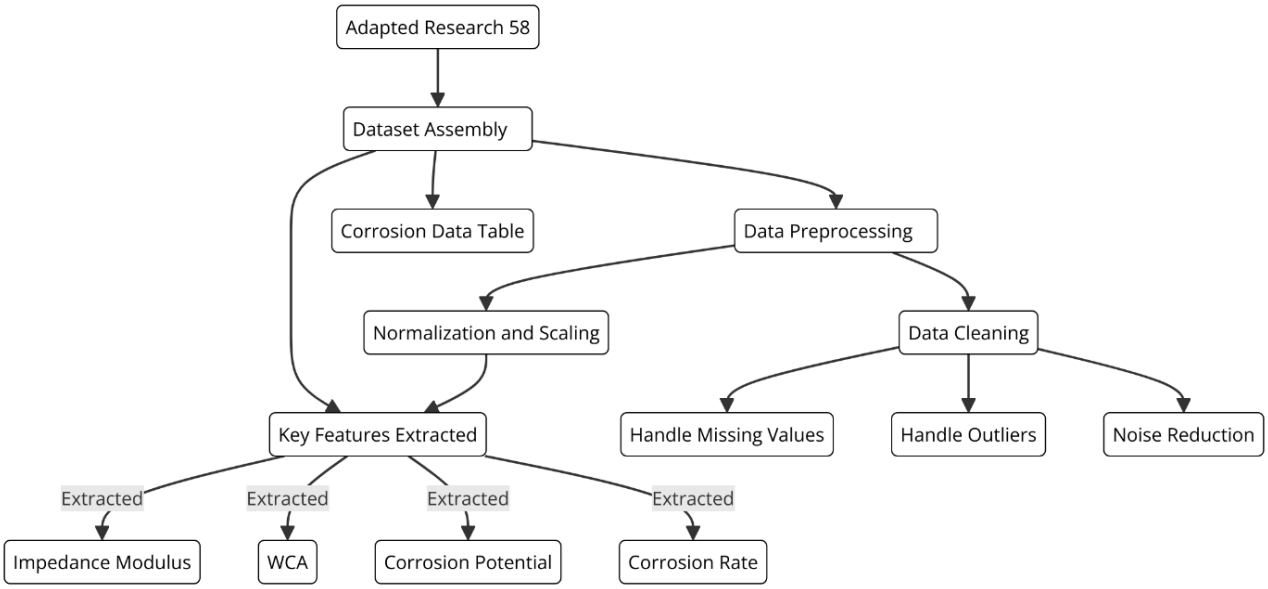

Fig. 1 Flowchart of data preprocessing and feature extraction for corrosion analysis

Table 1 Corrosion resistance and characterization of coatings

Sample	Nyquist Plot Impedance Modulus |Z| (Ω·cm²)	WCA (° ± SD)	Ecorr (V)	icorr (A/cm²)
ZE21B	193	45.6° ± 2.3	-1.701	8.418E-05
PMBP	45710	40.8° ± 1.5	-1.566	4.904E-07
MgF_2	49353	7.95° ± 0.4	-1.620	4.176E-07
SF	64122	28.5° ± 3.2	-1.549	3.659E-07
SF/(EGCG-Cu) ₂	70295	37.0° ± 6.6	-1.499	3.661E-07
SF/(EGCG-Cu) ₂/HA	93478	26.1° ± 2.1	-1.476	2.017E-07

model endothelial responses, offering insights that may optimize drug delivery and therapeutic interventions. This study addresses key challenges such as overfitting and data noise through advanced analytics, ensuring robust, interpretable predictions that could transform the understanding of endothelial function in cardiovascular health (Shariati *et al*. 2022).

2. Materials and methods

2.1 Data collection

In this study, the data collection process focused on assembling a high-quality dataset specifically targeting endothelial cell responses to CSV, using information adapted from the foundational research by Qianying Jia *et al*., titled "endothelium-mimicking with NO-generating coating on bioabsorbable magnesium alloy for improving corrosion resistance and biological responses of vascular stents (Jia *et al*. 2024). This dataset provided vital insights into endothelial cell adhesion strength under CSV. It included critical features such as impedance modulus (|Z|), WCA, Electrochemical corrosion potential (Ecorr), and corrosion rate (Shariati *et al*. 2020a). These parameters were chosen to reflect essential aspects of endothelial cell responses, precisely adhesion strength and corrosion resistance, directly aligning with the objectives of this research (Zhang *et al*. 2024). To ensure data accuracy and model compatibility, preprocessing steps were rigorously applied. Data cleaning methods addressed common issues such as missing values, outliers, and noise, typical in biological datasets. These measures were essential to enhance data reliability and minimize distortions affecting the model's generalization ability. Normalization and scaling techniques were applied to bring all features into comparable ranges, which is critical for compatibility with ML models sensitive to data scale variations (Song and Liu 2021). Fig. 1 illustrates the data preprocessing and feature extraction workflow for corrosion analysis, including steps such as dataset assembly, data cleaning, normalization, and extraction of key features like impedance modulus, WCA, corrosion potential, and corrosion rate. Table 1 illustrates the corrosion resistance and characterization of various coatings, providing a comparative analysis of their performance and properties.

Furthermore, data transformation and augmentation techniques were employed to improve the model's robustness and adaptability. These techniques expanded the dataset's diversity, enabling the capture of a broader spectrum of endothelial responses under varying CSV conditions. Simulating different experimental scenarios through transformations made the dataset more versatile, thus

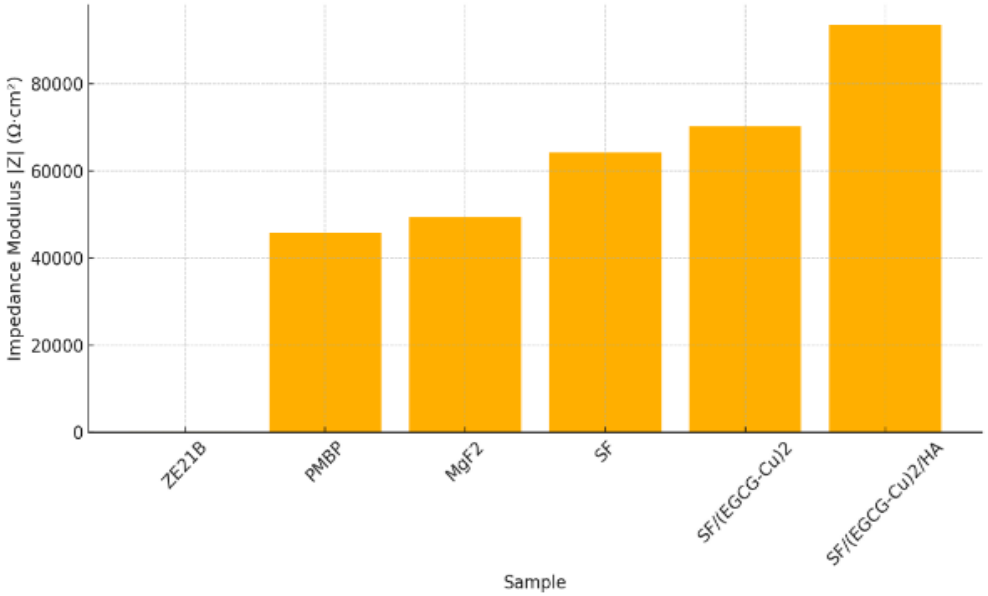

Fig. 2 Bar Chart Comparing Impedance Modulus (|Z|) Values Across Different Samples (ZE21B, PMBP, MgF2, SF, SF/(EGCG-Cu)2, SF/(EGCG-Cu)2/HA)

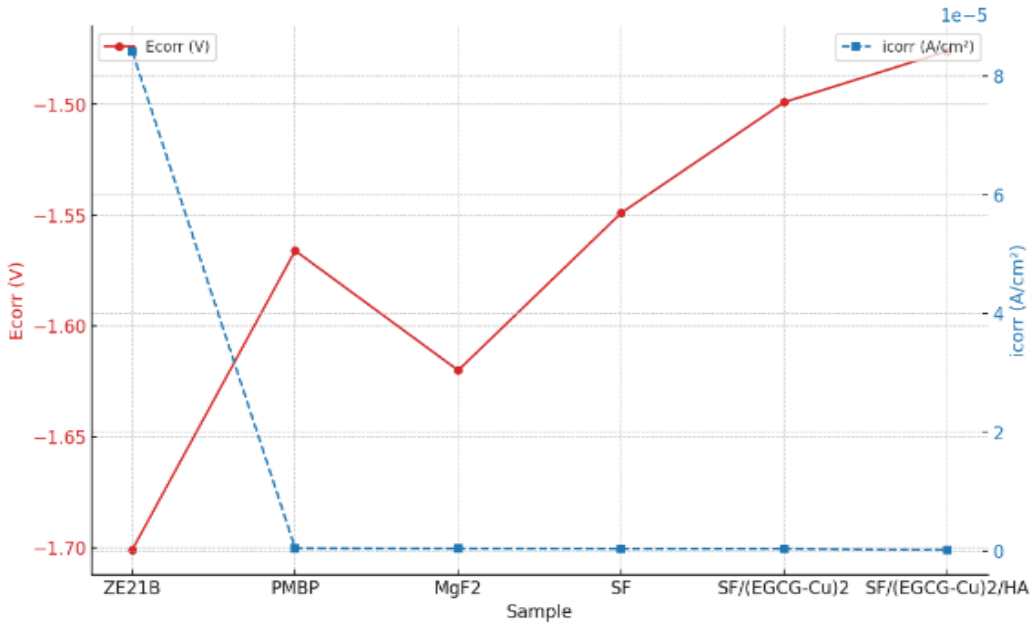

Fig. 3 Comparison of Ecorr (V) and Icorr (A/cm²) Across Samples ZE21B, PMBP, MgF2, SF, SF/(EGCG-CU)2, SF/(EGCG-CU)2/HA

enhancing the ML model's capacity to make accurate and generalizable predictions across varied contexts. Each step, from source selection to data cleaning, normalization, and augmentation, was essential in preparing a dataset that accurately represents endothelial cell behavior under CSV conditions and meets practical ML analysis requirements. Fig. 2 compares the impedance modulus $|Z|$ (in $\Omega \cdot \mathrm{cm}^2$) across various samples. The sample "SF/(EGCG-Cu)2/HA" exhibits the highest impedance, indicating superior resistance, while "ZE21B" has the lowest. This suggests improved electrochemical properties for the "SF/ $(\mathrm{EGCGCu})2/\mathrm{HA}^2$ sample compared to others. Fig. 3 compares the corrosion potential (E_{corr}) and corrosion current density (i_{corr}) across various samples, illustrating corrosion resistance. E_{corr} (red line) shows a progressive increase from ZE21B to SF/(EGCG-Cu)2/HA, indicating improved corrosion resistance. Meanwhile, i_{corr} (blue dashed line) remains nearly zero for most samples, reflecting minimal corrosion rates. The sample $\mathrm{SF}/(\mathrm{EGCG}-\mathrm{Cu})2/\mathrm{HA}$, with the highest E_{corr} and lowest i_{corr}, demonstrates the best corrosion resistance, highlighting its superior electrochemical properties.

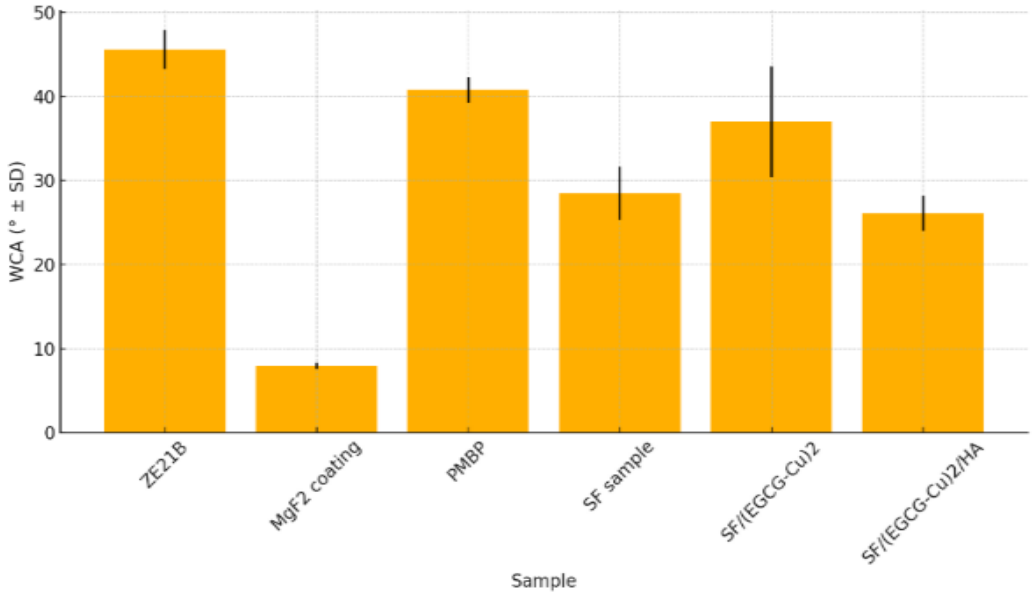

Fig. 4 WCA with SD for Samples ZE21B, MgF2 Coating, PMBP, SF Sample, SF/(EGCG-CU)2, and SF/(EGCG-CU)2/HA

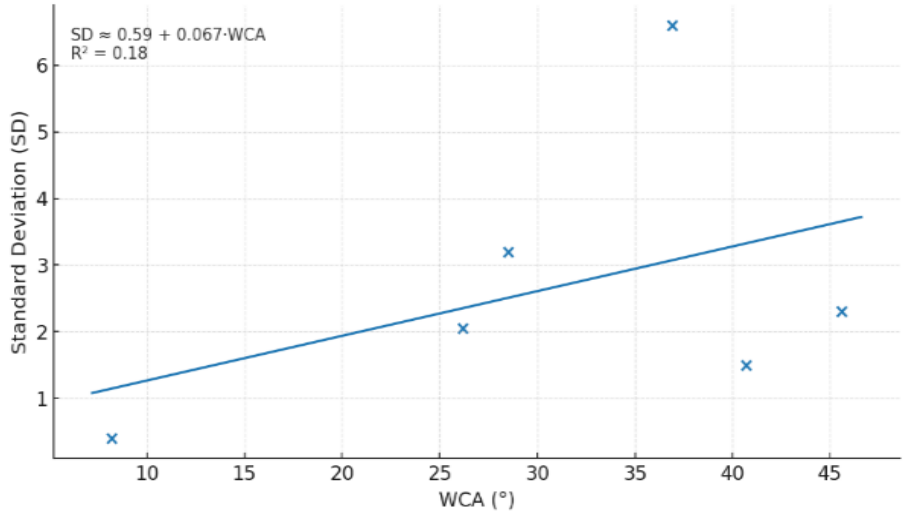

Fig. 5 Relationship between WCA and SD for Different Samples

Table 2 Coating Characterization - WCA

Sample	WCA
ZE21B	45.6° ± 2.3°
MgF_2 coating	7.95° ± 0.4°
PMBP	40.8° ± 1.5°
SF sample	28.5° ± 3.2°
SF/(EGCG-Cu) $_2$	37.0° ± 6.6°
SF/(EGCG-Cu) $_2$/HA	26.1° ± 2.1°

2.2 ML models

This study's selection of ML models was grounded in their ability to capture complex, non-linear patterns and interactions inherent to biological data. Neural networks, RF-fuzzy logic, and SVMs were chosen because of their strengths in addressing varied data structures and model complexity requirements. Neural networks are particularly adept at modeling non-linear relationships and capturing high-level abstractions in data, making them suitable for handling the intricacies of endothelial cell responses. RF-fuzzy logic provides robust interpretability and performs well on structured data using an ensemble of decision trees, capturing diverse feature interactions while minimizing overfitting risks (Kasahun and Legesse 2024). SVM, beneficial with high-dimensional datasets, is highly effective for classification tasks with a well-defined margin, making them appropriate for differentiating patterns in endothelial responses under different CSV conditions. Each model brings unique advantages, creating a comprehensive analytical approach to the study's objectives (Sanlı *et al.* 2020). Table 2 provides the coating characterization based on the WCA, demonstrating the hydrophobicity properties of the coated surfaces.

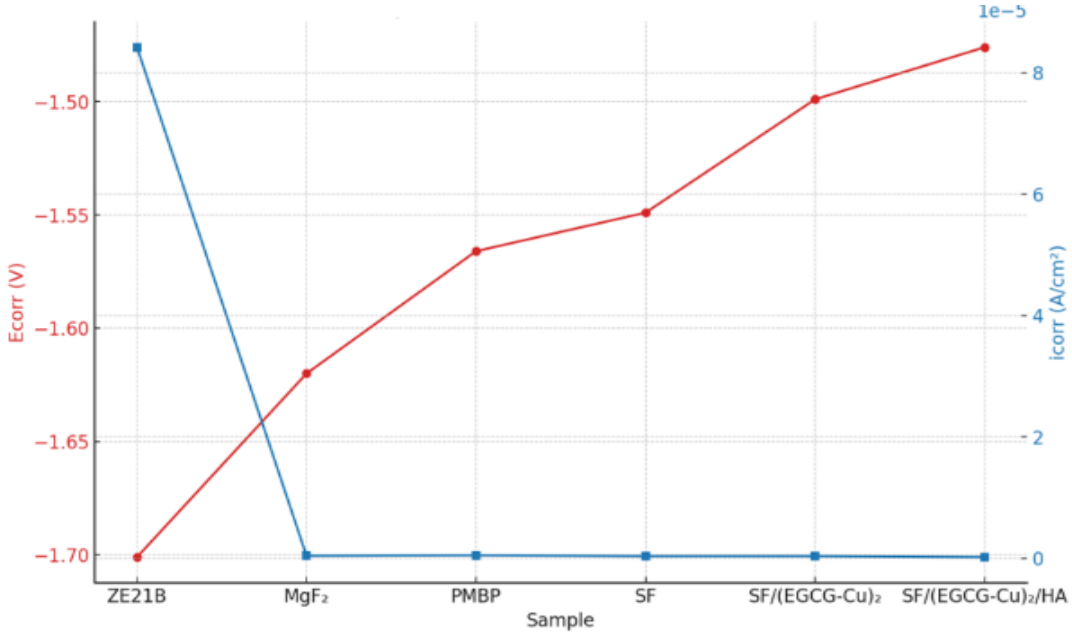

Fig. 6 Comparison of Ecorr (V) and icorr (A/cm²) values across different samples in PDP analysis

2.2.1 Neural networks

The architecture and structure of the neural network model were tailored to capture the complex patterns in endothelial cell responses under CSV conditions. The model architecture includes multiple layers, starting with an input layer that ingests data related to parameters like impedance modulus (|Z|), WCA, Ecorr, and corrosion rate. Hidden layers were configured with a variable number of neurons to allow sufficient capacity to learn the intricate relationships within the data (Sanlı *et al*. 2020). Activation functions such as Rectified Linear Unit (ReLU) were chosen for the hidden layers due to their efficiency in learning non-linear patterns while avoiding the vanishing gradient problem, which can hinder learning in deep networks. Fig. 4 displays the WCA for different samples, reflecting their surface wettability. ZE21B exhibits the highest WCA, indicating a more hydrophobic surface, while the MgF2 coating shows the lowest WCA, suggesting more excellent hydrophilicity. Other samples, such as PMBP, SF/(EGCG-Cu)2, and SF/(EGCG-Cu)2/HA, demonstrate intermediate values, implying moderate wettability. The error bars represent the Standard Deviation (SD), highlighting the consistency of the measurements. Fig. 5 illustrates the WCA and SD relationship for various samples. As WCA increases, the SD generally appears to vary, with some higher WCA values associated with larger SDs, indicating more significant variability in measurements. Samples with lower WCA exhibit smaller SDs, suggesting more consistent results for hydrophilic surfaces. This plot highlights the variability of surface wettability across different samples. The output layer utilized a sigmoid activation function, suitable for binary classification tasks, as it maps output probabilities between 0 and 1, enabling straightforward interpretation of endothelial cell responses (Park and Kim 2021). The configuration of the neural network was optimized through several key hyperparameters. The learning rate was set at a moderate level to balance convergence speed and model stability (Arevalo-Rodriguez *et al*. 2020). At the same time, the number of epochs was determined based on validation performance to avoid underfitting or overfitting (Zhou *et al*. 2024). Batch size was selected to control memory usage and influence the model's generalization ability across the dataset (Barkhordari and Qi 2025).

Hidden layers used regularization methods like dropout to prevent overfitting by training the network to build strong feature representations by arbitrarily setting a percentage of input units to zero. The goal of fine-tuning this dropout rate was to strike a compromise between the capacity of the model and its generalizability (Hebbel *et al*. 2020, Zhao *et al*. 2025). Fig. 6 illustrates the variation in Ecorr (V) (red line) and icorr (A/cm²) (blue line) across different samples. The Ecorr values increase progressively, indicating improved corrosion resistance, while icorr decreases significantly, highlighting reduced corrosion rates. This trend reflects better protective properties for the SF/(EGCG-Cu)₂/HA coating than other samples. Table 3 illustrates the Potentio Dynamic Polarization (PDP) parameters, showcasing the key electrochemical characteristics analyzed in the study. The dataset was partitioned into training, validation, and test sets during training so the model could learn and assess its performance without overfitting (Martí-Carvajal *et al*. 2015a). The data was divided as follows: 70% for training, 15% for validation, and 15% for testing, following a usual split ratio of 70:15:15 (Montero-Oleas *et al*. 2020, Mayorga-Ramos *et al*. 2022). The model's weights were fine-tuned using the Adam optimizer. This well-liked optimization method combines the strengths of RMSProp and the Adaptive Gradient Algorithm (AdaGrad) for more rapid and steady convergence (Simancas-Racines *et al*. 2015).

The Adam optimizer's capacity to dynamically modify

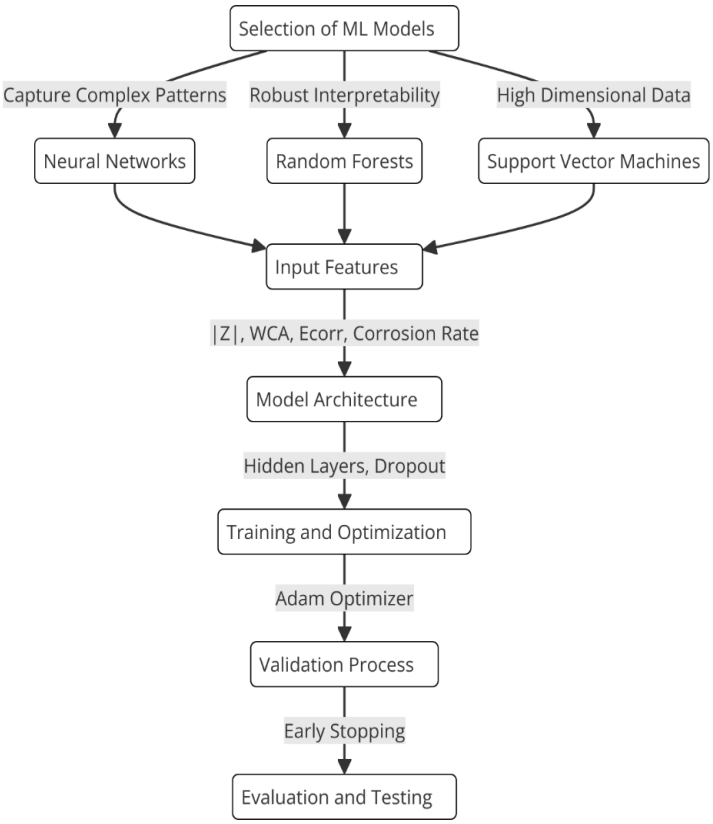

Fig. 7 Flowchart for machine learning model selection and development process

Table 3 PDP Parameters

Sample	Ecorr (V)	icorr (A/cm²)
ZE21B	-1.701 V	8.418E-05 A/cm²
MgF_2	-1.620 V	4.176E-07 A/cm²
PMBP	-1.566 V	4.904E-07 A/cm²
SF	-1.549 V	3.659E-07 A/cm²
SF/(EGCG-Cu) $_2$	-1.499 V	3.661E-07 A/cm²
SF/(EGCG-Cu) $_2$/HA	-1.476 V	2.017E-07 A/cm²

Table 4 Weight Loss and Corrosion Rate

Sample	Weight Loss after Immersion in SBF
ZE21B	Largest weight loss observed
SF-based samples	Minimum weight loss, indicating better resistance

learning rates depending on data properties was a deciding factor for complicated biological datasets with different feature distributions (Fuentes *et al*. 2018, Simancas-Racines *et al*. 2018). Various evaluation techniques were implemented during training to ensure the model's performance was generalized well (Núñez-González *et al*. 2020c). Cross-validation was used to validate model performance across different subsets of the training data, reducing the chances of model bias (Marti-Carvajal *et al*. 2015). Early stopping was incorporated to prevent overfitting; this involved halting training when the validation performance stopped improving, ensuring that the model did not memorize the training data but generalized well to unseen data (Molano-Franco *et al*. 2023, Zambrano *et al*. 2023). Additionally, loss and accuracy were continuously monitored during training to track the model's learning progression and fine-tune hyperparameters if needed (Franco *et al*. 2018b). This rigorous evaluation approach facilitated the development of a robust neural network model capable of reliably predicting endothelial cell adhesion strength and other responses under CSV conditions (Chérrez-Ojeda *et al*. 2018, Núñez-González *et al*. 2020b). Figure 7 illustrates the step-by-step flowchart for selecting and developing a machine learning model, outlining the process's key phases and decision points. Table 4 illustrates the relationship between weight loss and the corrosion rate, providing detailed measurements and trends observed in the study.

2.2.2 RF with fuzzy integration

The RF model was implemented to leverage its ensemble-based approach, which combines multiple decision trees to improve predictive accuracy and robustness (Cherrez-Ojeda *et al*. 2020, Guevara-Ramírez *et al*. 2022). For this study, the configuration of the RF model included integrating fuzzy logic with a carefully selected number of trees (estimators) to ensure sufficient model stability while avoiding excessive computational demands (Kim 2024). Typically, between 100 and 500 trees were used, with an exact number chosen based on validation results to balance model complexity and performance (Simancas-Racines *et al*. 2023, Chapela *et al*. 2024). The maximum depth of each tree was also tuned; a controlled depth ensured that each

tree could capture necessary details in the endothelial response data without overfitting to noise (Guevara-Ramírez *et al*. 2023). Node splitting within the trees was based on criteria such as Gini impurity or entropy, with Gini impurity preferred for its efficiency in assessing split quality (Bao *et al*. 2018, Chen *et al*. 2025). The integration of fuzzy logic introduced an additional layer of decision-making at the split level, where fuzzy membership functions were used to evaluate the degree of data points belonging to specific categories (Arevalo-Rodriguez *et al*. 2021, Ruiz-Pozo *et al*. 2023). This approach enabled the RF model to make more nuanced splits at each node, improving its ability to classify diverse endothelial responses under CSV conditions. By incorporating fuzzy logic, the RF model gained an enhanced capacity to manage the variability and ambiguity inherent in biological data, further improving its predictive performance and reliability in analyzing endothelial adhesion strength (Li *et al*. 2024, Liu *et al*. 2024). Hyperparameter tuning was critical in optimizing the RF model for reliable and generalized performance (Martí-Carvajal *et al*. 2015b). A combination of grid search and cross-validation techniques was employed to refine parameters like tree depth, the number of trees, and the minimum samples required to split a node (Montero-Oleas *et al*. 2017, Simancas-Racines *et al*. 2019). Grid search allowed for systematic exploration of hyperparameter combinations, while cross-validation evaluated model performance across different data subsets, reducing the likelihood of overfitting (Zandi *et al*. 2018, Shariati *et al*. 2020b). This tuning helped optimize the model's accuracy and precision, capturing meaningful patterns within the complex endothelial cell data (Ghezelbash *et al*. 2019, Daviran *et al*. 2021). An essential feature of the RF model in this study was its ability to provide feature importance scores, which quantified the contribution of each input feature (e.g., impedance modulus |Z|, WCA, Ecorr, and corrosion rate) to the model's predictions (Paz-Cruz *et al*. 2023, Ruiz-Pozo *et al*. 2024). By analyzing the importance of features, the model highlighted which parameters had the most substantial impact on endothelial cell adhesion strength (Trung *et al*. 2019). This interpretative aspect of RF was precious, as it offered insights into the underlying biological factors influencing endothelial responses under CSV conditions (Montesinos-Guevara *et al*. 2022, Zambrano *et al*. 2024b). Understanding the relative importance of features helped clarify each parameter's mechanistic role, guiding the interpretation of model results and identifying which characteristics are most critical in endothelial cell behavior (Bonfill *et al*. 2015). This interpretability aligns with the study's objectives to predict responses and deepen understanding of CSV's mechanistic effects on endothelial function (Sun *et al*. 2019).

2.2.3 SVMs

The SVM model was chosen for its effectiveness in handling high-dimensional data and its capability to distinguish complex, non-linear patterns, making it suitable for predicting endothelial cell responses under CSV (Simancas-Racines *et al*. 2017). Various kernel types were explored during model development, including linear and Radial Basis Function (RBF) kernels. The RBF kernel was ultimately selected as the most suitable due to its flexibility in capturing non-linear relationships among features like impedance modulus (|Z|), WCA, Ecorr, and corrosion rate (Zambrano-Sánchez *et al*. 2022). Unlike the linear kernel, which assumes linear separability, the RBF kernel can map the input data into a higher-dimensional space, effectively capturing the intricate dependencies between features that characterize endothelial cell responses (Chapela *et al*. 2023, Yang *et al*. 2025). Parameter optimization was crucial to developing the SVM model to ensure balanced performance and robustness (Chérrez-Ojeda *et al*. 2020). The critical parameters optimized were the regularization parameter (C) and the kernel coefficient (gamma). The regularization parameter C controls the trade-off between achieving a low error on the training data and maintaining a smooth decision boundary, which prevents the model from overfitting (Salazar *et al*. 2021, Muscogiuri *et al*. 2024). A high C value emphasizes the correct classification of training samples, while a lower value allows for a more generalized decision boundary (Verde *et al*. 2024). The gamma parameter, specific to the RBF kernel, defines the influence of individual training examples (Song *et al*. 2024, Peng *et al*. 2025). Higher gamma values mean that points closer to the decision boundary are given greater importance, allowing the model to capture complex patterns but at the risk of overfitting (Simancas-Racines *et al*. 2019, Reytor-González *et al*. 2024). A grid search method was used to systematically explore different combinations of C and gamma values, with cross-validation applied to identify the settings that achieved the best accuracy without overfitting (Shao *et al*. 2012). The training process for SVMs involved splitting the dataset into training, validation, and test sets to evaluate model performance across different phases (Jahandari *et al*. 2022). This step ensured that the model was accurate on the training data and generalized well to new, unseen data (Cadena-Ullauri *et al*. 2024b, Tamayo-Trujillo *et al*. 2024). In cases where data imbalances were present for example if specific endothelial responses were underrepresented resampling techniques or class-weight adjustments were applied to balance the influence of each class during training (Yang *et al*. 2022, Yang *et al*. 2024). This approach helped the SVM model remain sensitive to all categories of endothelial responses, improving its reliability across a diverse range of CSV conditions (Calderón *et al*. 2017). By carefully tuning parameters and managing data imbalances, the SVM model was optimized to deliver high accuracy and meaningful predictions for endothelial adhesion strength and other response metrics, aligning with the study's goals of using ML to deepen understanding of endothelial function under CSV (Yang *et al*. 2023, Zambrano *et al*. 2024a). Table 5 illustrates the results of cell viability and NO release, providing detailed insights into the cellular responses under different experimental conditions.

2.3 Performance metrics

The performance of ML models was evaluated using

Table 5 Cell Viability and NO Release

Sample	NO Release Level	Contribution to Endothelial Function
SF/(EGCG-Cu) $_2$	High NO release	Improves endothelial function
SF/(EGCG-Cu) $_2$/HA	High NO release	Improves endothelial function

both classification and regression metrics (Núñez-González *et al*. 2020d). For classification, accuracy was defined as $(TP + TN)/(TP + TN + FP + FN)$, providing the overall success rate of predictions (Hidalgo *et al*. 2012, AbuShanab *et al*. 2023). To account for class imbalances, precision ($TP/(TP + FP)$) measured the reliability of positive predictions, while recall ($TP/(TP + FN)$) assessed the ability to capture all true positives. Their balance was expressed by the F1 score, calculated as $2 \times$ (Precision $\times$ Recall)/(Precision + Recall) (Rubio-Gallegos *et al*. 2019). For regression tasks such as predicting impedance modulus values, errorbased metrics were used (Choi *et al*. 2023, Abood *et al*. 2024): Mean Absolute Error (MAE), given by $MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$, representing the average prediction error (Núñez-González *et al*. 2020a), and Mean Squared Error (MSE), defined as $MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$, which emphasized larger deviations (Franco *et al*. 2018a, Cadena-Ullauri *et al*. 2024a).

For models like RF, feature importance analysis was performed on predictions to determine the influence of individual features, such as impedance modulus, water contact angle, and Ecorr (Li *et al*. 2019). (Robeson and Willmott 2023). Traditional feature importance was calculated using the mean decrease in impurity or Gini importance across decision trees, with higher scores indicating greater relevance in predicting endothelial responses. To enhance this analysis, fuzzy logic was integrated, allowing the assessment of feature contributions on a continuum rather than as binary outcomes (Mohammadhassani *et al*. 2013). Fuzzy membership functions assigned degrees of importance to each feature, accommodating uncertainty and imprecision in the biological data. This integration enabled a more nuanced understanding of the critical factors influencing endothelial adhesion strength and responses to CSV, providing a robust framework for analyzing complex biological interactions (Morales *et al*. 2021, Paz-Cruz *et al*. 2024). To further support model interpretability, visualization methods, such as Shapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME), were utilized (Avila *et al*. 2024). These methods helped in explaining individual model predictions, thereby bridging the gap between ML output and actionable biological insights. SHAP values attributed portions of the prediction outcome to specific features (Viteri-García *et al*. 2018, Montesinos-Guevara and Simancas-Racines 2019). At the same time, LIME provided locally interpretable models around each prediction, allowing more precise translation of complex ML findings into meaningful interpretations for endothelial cell function analysis (Canbek 2023).

2.4 Data processing workflow

The data processing workflow was designed to ensure that the ML models were trained, validated, and tested systematically and unbiasedly, maximizing their predictive accuracy while minimizing overfitting (Verdugo-Paiva *et al*. 2020). Data Splitting was conducted to allocate portions of the dataset for different stages of model development. The dataset was divided into training, validation, and testing (Nugroho *et al*. 2023). The training set was used to develop and adjust the model's parameters, while the validation set provided a basis for tuning model hyperparameters and evaluating performance mid-process. Finally, the testing set assessed the model's effectiveness on unseen data, simulating real-world conditions and enabling an unbiased evaluation of generalization capabilities (López *et al*. 2023). Typically, an 80-10-10 or 70-15-15 split was applied, ensuring a balanced distribution across all subsets and preserving the integrity of the data throughout training and evaluation. Cross-validation was implemented to enhance the model's robustness, particularly as a preventive measure against overfitting. Using this method, the training data was folded into many sets; one set was used as a validation set, while the other was used as the training set (Shariati *et al*. 2021a). For example, a 5-fold cross-validation involves dividing the dataset into five sections and training the model five times, with one section used for validation and the other for training (Safa *et al*. 2020). Evaluating the model over multiple iterations ensures that a small subset does not influence performance results, providing a more accurate representation of the model's correctness (Shariati *et al*. 2020a). Cross-validation also proved helpful in models where the dataset size might be limited, as it maximized data usage by enabling each data point to serve in both training and validation roles (Parvandeh *et al*. 2020, Shen *et al*. 2020). Overfitting management strategies were crucial to ensure that the models did not simply memorize the training data but learned to generalize across new, unseen data (Armaghani *et al*. 2020). Several techniques were employed depending on the model type (Toghroli *et al*. 2014). Regularization techniques like L2 were applied in neural networks, penalizing large weights to discourage overly complex models that might fit noise in the data. Additionally, early stopping was used to halt training once the validation performance plateaued or began to decline, preventing the model from overfitting on the training set (Kernbach and Staartjes 2022, Tran *et al*. 2022). Pruning and hyperparameter tuning were used for RF models to reduce model complexity and avoid overfitting. Pruning limited the depth and size of individual trees within the forest, preventing them from growing overly specific to the training data, which could impair generalizability (Sedghi *et al*. 2018, Shariati *et al*. 2022). To further enhance this process, fuzzy logic was integrated to manage the uncertainty in hyperparameter selection and pruning criteria. For instance, fuzzy membership functions assessed the trade-offs between model complexity and performance, enabling more flexible and nuanced decision-making during pruning (Shariati *et al*. 2021b). Hyperparameters such as the number of trees, maximum tree depth, and minimum

samples per leaf were adjusted using grid search or cross-validation techniques, incorporating fuzzy-based evaluations to identify configurations that minimized overfitting while accounting for imprecisions in data patterns (Katebi *et al*. 2020, Yazdani *et al*. 2021b).

3. Results and discussion

3.1 Model performance

A structured approach is essential to comprehensively compare and analyze the ML models, beginning with clearly defining evaluation metrics. This study's metrics are chosen based on their relevance to assessing endothelial cell responses under CSV. For classification, metrics such as accuracy, precision, recall, and F1 score will provide insights into each model's ability to classify and balance predictions correctly (Kernbach and Staartjes 2022). Accuracy reflects the overall success rate, while precision and recall help gauge the proportion of correct optimistic predictions and the model's sensitivity to actual positive cases, respectively. The F1 score combines these into a single measure, which is especially valuable for handling imbalanced data. If continuous values are predicted, like impedance modulus, error metrics like MAE and MSE will capture the average prediction error and emphasize more significant deviations. Additionally, interpretability metrics, such as feature importance from RF–fuzzy and tools like SHAP and LIME, will clarify the biological factors that influence model predictions. Next, data is split into training, validation, and testing subsets to ensure unbiased model evaluation, typically in a 70-15-15 or 80-10-10 ratio. Cross-validation, particularly k-fold cross-validation is applied to mitigate overfitting by assessing model performance across various subsets of the training data. This process improves the reliability of performance estimates and maximizes data utilization, which is especially beneficial when data is limited. In the training phase, hyperparameter tuning for each model enhances performance and ensures robustness. For neural networks, settings like learning rate, batch size, and dropout rate are tuned using grid or random search, with validation performance monitored to prevent overfitting. RF-fuzzy parameters, such as the number of trees, depth, and split criteria, are optimized to balance complexity and accuracy. SVM tuning focuses on kernel type, regularization (C), and kernel coefficient (gamma), with grid search and cross-validation ensuring each model can generalize well. Once trained, the models are evaluated on the testing set using predefined metrics. Visualization techniques, including AUC-ROC and precision-recall curves, assist in comparing classification models, while scatter plots of predicted versus actual values provide insight into regression model performance. Further, feature importance analysis and interpretability tools, such as SHAP, are employed to explore which parameters, impedance modulus or Ecorr, are most influential in predictions. For a deeper understanding, error analysis identifies cases where models underperform, potentially revealing data patterns or feature relationships that influence errors. Overfitting checks involve comparing training and validation scores; a significant disparity suggests additional regularization might be necessary. Interpretability tools also enhance biological insights by clarifying model predictions, with SHAP values showing individual feature contributions. Finally, findings are summarized, emphasizing each model's strengths in prediction and interpretability. Statistical tests, like paired t-tests or Wilcoxon signed-rank tests, determine if performance differences are statistically significant. Balancing quantitative metrics with interpretability, this robust evaluation approach provides insights into the models' predictive power and underlying biological relationships, aligning with the study's objectives of understanding endothelial responses to CSV conditions. This careful integration ensures the findings are both actionable and scientifically robust, laying the foundation for future applications in clinical or therapeutic contexts.

3.2 Model performance evaluation

Three ML models, neural networks, RF-fuzzy, and SVM, were tested for their ability to predict endothelial cell responses under CSV. Various metrics were utilized for evaluation, including accuracy, precision, recall, F1 score, MAE, and MSE. The model's capacity to categorize, forecast, and generalize the biological consequences associated with endothelial cell adhesion under different situations is enhanced by each measure.

3.2.1 Accuracy

Accuracy reflects the overall correctness of a model's predictions. The ratio of correct predictions to the total number of forecasts determines it.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

where TP = True Positives, TN = True Negatives, FP = False Positives, and FN = False Negatives. The provided confusion matrix can be utilized to compute the accuracy of each model, enabling a comparison of their overall performance. For instance, if the confusion matrix for one model shows the following values:

$$TP = 50,\ TN = 45,\ FP = 5,\ FN = 10 \tag{2}$$

Then,

$$\text{Accuracy} = \frac{50 + 45}{50 + 45 + 5 + 10} = \frac{95}{110} = 0.8636 \text{ or } 86.36\% \tag{3}$$

3.2.2 Precision

Precision shows how many of the predicted positive cases were positive. It is beneficial when minimizing false positives, which is crucial for predicting critical endothelial responses such as

$$\text{Precision} = \frac{TP}{TP + FP} \tag{4}$$

$$\text{Precision} = \frac{50}{50 + 5} = \frac{50}{55} = 0.9091 \text{ or } 90.91\% \tag{5}$$

3.2.3 Recall (Sensitivity)

Recall or sensitivity or the true positive rate indicates the proportion of actual positive cases the model successfully identifies. This metric is essential when capturing all positive cases, such as detecting endothelial dysfunction as:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{6}$$

$$\text{Recall} = \frac{50}{50 + 10} = \frac{50}{60} = 0.8333 \text{ or } 83.33\% \tag{7}$$

The F1 score takes the harmonic mean of recall and accuracy to achieve a balanced trade-off medium between the two measures. Because it treats false positives and false negatives equally, it works well with datasets that aren't balanced:

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{8}$$

If a model has a precision of 0.67 and a recall of 0.625, the F1 score would be:

$$\text{F1 Score} = 2 \times \frac{0.67 \times 0.625}{0.67 + 0.625} = 0.647 \text{ or } 64.7\% \tag{9}$$

The F1 score is a valuable metric, providing a single statistic that integrates both recall and precision. Because it treats false positives and false negatives equally, it is particularly effective for imbalanced datasets and is especially relevant for predicting endothelial adhesion strength, where it helps improve the reliability of positive predictions while reducing the likelihood of missing essential adhesion cases.

3.2.4 Confusion matrix

A confusion matrix provides a summary of prediction results for classification problems. It displays the number of correct and incorrect predictions made by the model, helping visualize the performance across all classes. The confusion matrix is often used as the basis for computing the metrics mentioned above.

3.3 MAE and MSE

For regression tasks, where the model predicts continuous values like impedance modulus (|Z|), MAE and MSE are used to measure the prediction errors. The MAE is the mean of the absolute differences between predicted and actual values, reflecting the average magnitude of errors in a set of predictions as:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{10}$$

where y_i is the actual value, $\hat{y}_i$ is the predicted value, and n is the number of predictions.

For example, if the true values for a sample of 5 observations are [3,5,7,9,11] and the predicted values are [2.8, 5.1, 6.9. 8.7, 10.8], the MAE is:

$$\text{MAE} = \frac{1}{5}(|3 - 2.8| + |5 - 5.1| + |7 - 6.9| + |9 - 8.7| + |11 - 10.8|) = \frac{1}{5}(0.2 + 0.1 + 0.1 + 0.3 + 0.2) = \frac{0.9}{5} = 0.18 \tag{11}$$

3.3.1 MSE

The MSE measures the average of the squared differences between predicted and actual values, giving more weight to larger errors as

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{12}$$

$$\text{MSE} = \frac{1}{5}((3 - 2.8)^2 + (5 - 5.1)^2 + (7 - 6.9)^2 + (9 - 8.7)^2 + (11 - 10.8)^2) \tag{13}$$

This value indicates that the model's predictions deviate on average by $\sqrt{0.038} \approx 0.194$ units from the actual values.

3.4 Data splitting and cross-validation

The preparation of datasets for ML models relies heavily on data splitting. This procedure guarantees that models are trained, verified, and tested on separate, non-overlapping subsets of data. This approach allows for assessing whether the model has overfitted to the training data and evaluating its generalizability. Training Set was one of three primary subsets of the dataset employed in this research. The model takes this data, learns the patterns, and then tweaks its internal parameters to get the best match. It usually takes up most of the dataset since the model needs a lot of data to find patterns. A validation set is used to assess the model's performance during training and to fine-tune its hyperparameters. This is made possible by optimizing the model (by adjusting the learning rate or applying regularization, for example) and receiving feedback on how well it will perform on unknown data. Once the model has been trained and refined, it is time to evaluate its performance using the test set. To get an accurate picture of the model's performance in the actual world, this subset is crucial for determining its generalizability to fresh, unknown data. Data was typically divided as follows: 70% for training, 15% for validation, 15% for testing, 80% for training, 10% for validation, and 10% for testing, according to this research. The dataset was partitioned into training, validation, and testing subsets to ensure unbiased evaluation of model performance. A 70-15-15 split ratio is frequently adopted in studies where greater emphasis is placed on model validation and testing, thereby reducing the risk of bias and overfitting. Alternatively, an 80-10-10 split ratio is commonly applied when sufficient data are available, allowing a larger portion to be used for training and improving the model's learning capacity. The choice of split ratio depends on dataset size and research objectives, with smaller datasets benefiting from proportionally larger validation and test sets to preserve generalizability.

3.5 Cross-validation

Cross-validation is a reliable method for evaluating the

model's efficacy while minimizing the possibility of overfitting. By dividing the data into smaller sets, or "folds," cross-validation allows the model to be trained and evaluated on numerous combinations of these folds rather than just one set. In k-fold cross-validation, the dataset is split into k equal-sized folds, usually 5 or 10 folds, which is the most used method. By dividing the dataset into k subsets or folds, k-fold cross-validation may be achieved. After training on k-1 folds, the model is tested on the last fold. This process is repeated k times, alternating the test and training sets to ensure each iteration utilizes a different fold. The mean score from all k-test sets is the ultimate measure of performance. One example is a 5-fold cross-validation, which involves dividing the data into five different groups as the model is trained on the first four folds in the first iteration and tested on the 5th, in the second iteration, the model is trained on folds 1, 2, 3, and 5 and tested on fold 4 and this continues until each fold is used once as the test set.

Cross-validation tests the model on several data groups rather than just one to reduce the likelihood of overfitting. Because it considers the possibility of data fluctuation, this technique gives a more reliable and precise assessment of the model's performance. It also maximizes the use of available data, especially in cases where the dataset is limited. By performing k-fold cross-validation, variance in performance estimation can be reduced. Cross-validation ensures that the model's performance is consistent across different subsets of data, minimizing the likelihood of a high variance that could arise from a single train-test split. This approach enhances model robustness by training and testing on various data subsets, ensuring the model fits a specific portion of the dataset and adapts to diverse data points while maximizing data utilization. In situations where the dataset size is limited, each data point is used for training and validation, improving the efficiency of data use.

3.6 Hyperparameter tuning

Hyperparameter tuning is crucial in optimizing ML models by adjusting model parameters to improve their predictive power. Hyperparameter tuning in neural networks and several key hyperparameters influence the model's performance. The learning rate (η) is a critical parameter that controls how much the model's weights are adjusted during each training step. The Eq used to update the weights in gradient descent is:

$$\theta = \theta - \eta \cdot \nabla_\theta J(\theta) \tag{14}$$

where θ represents the model's weights, η is the learning rate, and $\nabla_\theta J(\theta)$ is the gradient of the loss function concerning the weights.

This research uses several learning rates from training data to maximize the model's generalization capacity, such as 10^{-6} and 10^{-3}. Although quicker convergence is possible with higher learning rates, there is a danger of exceeding the minimal loss if this happens. Although slower rates enhance stability, convergence occurs at a reduced pace. A critical hyperparameter follows the batch size. It determines the number of samples that must be processed before updating the model's weights. Batch processing requires a modification to the gradient descent equation:

$$\theta = \theta - \frac{\eta}{B}\sum_{i=1}^{B} \nabla_\theta J\left(\theta^{(i)}\right) \tag{15}$$

where B is the batch size, $\nabla_\theta J\left(\theta^{(i)}\right)$ is the gradient of the loss function for each sample.

The optimal batch size typically ranges from 32 to 128. This study performs a grid search to find the batch size that minimizes the MAE or MSE (discussed later in performance metrics). The number of epochs controls how often the model is trained on the full dataset. This hyperparameter is fine-tuned by monitoring validation performance to avoid overfitting. The dropout rate (p) helps regularize the model by randomly setting a fraction p of input units to zero. Dropout is applied during training to prevent the model from memorizing the training data. In Random Forests (RF), several hyperparameters critically influence model performance. The number of trees (n_{trees}) determines the size of the ensemble; increasing the number of trees generally improves predictive accuracy by reducing variance but also raises computational cost. In practice, values between 100 and 500 are commonly evaluated to balance accuracy and efficiency. The maximum tree depth ($d_{\max}$) controls how many levels each tree may grow. Deeper trees can capture more complex relationships within the data but risk overfitting, whereas shallower trees may underfit. Optimal depth is typically identified through grid search combined with cross-validation. In the fuzzy-enhanced RF framework, fuzzy membership functions are applied to evaluate the trade-off between model complexity and generalization capacity, assigning weighted scores to both overfitting risk and interaction modeling ability. Additionally, the node splitting criterion is a key parameter, with Gini impurity or entropy commonly employed. For a dataset D with K classes, the Gini impurity is defined as:

$$\text{Gini}\,(D) = 1 - \sum_{k=1}^{K} p_k^2 \tag{16}$$

where p_k denotes the proportion of samples belonging to class k. This criterion guides the selection of feature splits, ensuring that each division maximally reduces class impurity and thereby improves predictive performance.

The minimum samples per split (m_{split}) parameter specifies the minimum number of samples required to split an internal node. Larger values reduce the likelihood of overfitting by ensuring that splits occur only when sufficient data are available. This parameter, along with others, is typically optimized using grid search combined with cross-validation, evaluating performance metrics such as accuracy, precision, recall, and F1 score. For Support Vector Machines (SVMs), the most critical hyperparameters include the kernel type, the regularization parameter (C), and the kernel coefficient (γ). The Radial Basis Function (RBF) kernel is frequently employed due to its capacity to capture nonlinear relationships, defined as:

$$K(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right) \tag{17}$$

where x and x' are data points, and σ controls the kernel width. The regularization parameter C governs the trade-off between minimizing training errors and maintaining a smooth decision boundary. Higher values of C prioritize correct classification of training data but may increase overfitting. The SVM decision function is expressed as:

$$f(x) = \sum_{i=1}^{n} \alpha_i y_i K(x_i, x) + b \tag{18}$$

where α_i are Lagrange multipliers, y_i the support vector labels, and b the bias term. The gamma parameter $\left(\gamma = 1/(2\sigma^2)\right)$ determines the influence of each training example: higher values increase sensitivity to nearby points but risk overfitting. Optimal values of C and γ are typically identified via grid search and cross-validation to balance bias-variance trade-offs. To ensure reliable performance under cyclic shear viscosity (CSV) conditions, hyperparameters in all models neural networks, RF-fuzzy, and SVM were tuned using systematic search strategies with cross-validation (e.g., 5-fold or 10-fold). Classification tasks were evaluated using accuracy, precision, recall, and F1 score, while regression tasks employed MAE and MSE. These complementary metrics provided a robust framework for assessing predictive performance and generalizability to unseen endothelial response data.

3.7 Feature importance and interpretability

This study employed the RF model to assess the feature importance of parameters related to endothelial cell behavior under CSV. The critical features identified for the model include impedance modulus (|Z|), WCA, Ecorr, and Corrosion Current Density (icorr). RF uses an ensemble of decision trees to make predictions, and one of the main advantages is the ability to calculate feature importance. This is done by assessing the decrease in Gini impurity as each feature is used to split the data. The decline in impurity is averaged across all trees in the forest, which allows the model to rank features based on their contribution to improving prediction accuracy. To enhance this process, fuzzy logic was incorporated to address the inherent uncertainties and imprecisions in the biological data. Fuzzy membership functions were applied to evaluate each feature's influence on the model's performance rather than assigning strict binary importance scores. For instance, impedance modulus (|Z|) and WCA, directly related to material adhesion and surface properties, were assigned fuzzy scores based on their graded relevance to endothelial cell responses under CSV conditions. In computing feature importance within the RF framework, the Gini index was calculated for each node and feature, with fuzzy logic modifying the evaluation to accommodate overlapping feature effects and uncertain data distributions. This fuzzy-enhanced approach provided a more nuanced understanding of the factors influencing endothelial responses, ensuring that critical features were accurately identified even under imprecise or variable conditions. Eq. (19) for Gini impurity remains central but is augmented with fuzzy weights to reflect feature relevance dynamically in uncertain contexts.

$$\text{Gini}_{\text{split}} = 1 - \sum_{i=1}^{K} p_i^2 \tag{19}$$

where p_i is the proportion of samples in class i at node, and K is the number of classes. After training the model, the importance scores are obtained by summing the decrease in Gini impurity for each feature across all trees in the forest. These scores are then used to rank the features by their predictive power. Following this, SHAP was used to understand the contribution of each feature to individual predictions. SHAP values, derived from cooperative game theory, quantify how much each feature contributes to the prediction by averaging the difference in model outputs for every possible combination of features. The Shapley value for each feature i is calculated as:

$$\phi_i = \frac{1}{|N|} \sum_{S \subseteq N \setminus [i]} [f(S \cup [i]) - f(S)] \tag{20}$$

where ϕ_i is the Shapley value for feature i, and $f(S)$ is the model's prediction when a set of features S is used. This method helps reveal whether increasing values of impedance modulus (|Z|) or WCA lead to higher adhesion strength and clarifies how these features impact predictions. In addition to SHAP, LIME was used to provide local interpretability. LIME works by generating a surrogate interpretable model, like linear regression, around each prediction and explaining how changes in feature values affect the prediction. For example, by perturbing the WCA and analyzing its effect on adhesion strength predictions, LIME can help quantify the role of this feature in determining endothelial cell behavior. In the context of the given data, impedance modulus (|Z|) and WCA are expected to have the highest scores of feature importance. Impedance modulus (|Z|) is directly related to the material's resistance to deformation under cyclic shear, which influences endothelial cell adhesion. WCA affects the wettability and surface energy of the material, playing a crucial role in determining endothelial cell attachment. Both features are critical in predicting how ECs respond to the surfaces exposed to CSV conditions. Using SHAP and LIME helps explain individual predictions and improves the interpretability of complex models like RF. These tools help bridge the gap between ML results and actionable biological understanding by offering insights into which features are most influential in predicting endothelial responses. The ability to explore the influence of features such as *Ecorr* or *icorr* on endothelial behavior is provided, offering insights into the mechanical effects of material properties on vascular health.

3.8 Error analysis and overfitting

Model performance was assessed using error-based metrics, specifically Mean Absolute Error (MAE) and Mean Squared Error (MSE). MAE quantifies the average magnitude of prediction errors and is defined as:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{21}$$

where y_i represents the actual value and $\hat{y}_i$ the predicted value. To emphasize larger deviations, MSE was also calculated as:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (22)$$

These errors were computed for each model (neural networks, RF-fuzzy, and SVM) on the test set to identify performance limitations. Overfitting was detected by comparing results across training, validation, and test sets. A substantial gap between training and validation accuracy indicated memorization of training data rather than generalizable learning (e.g., 90% training accuracy versus 70% validation accuracy). To mitigate overfitting, techniques such as L2 regularization, early stopping, and cross-validation were employed. 5-fold cross validation was applied, ensuring that each model was trained and validated on multiple data partitions. This procedure reduced variance in performance estimates and provided a more robust evaluation of generalizability. Fluctuations in performance across folds were also monitored as indicators of sensitivity to dataset variation, further informing overfitting risk.

3.9 Statistical significance and comparative analysis

Statistical significance and comparative analysis to determine if there are significant differences between the performance of the three machine learning models (neural networks, RF-fuzzy, and SVM) in predicting endothelial cell adhesion strength under CSV conditions, statistical tests were performed. A paired t-test was conducted to compare the performance metrics, specifically accuracy, between the models on the test set. This test assesses whether the means of two models' performance metrics are statistically different, helping to identify if one model consistently outperforms another. The null hypothesis (H_a) in this case was that there is no difference in performance, while the alternative hypothesis (H_1) was that there was a significant difference in performance as

$$t = \frac{\bar{d}}{\frac{s_4}{\sqrt{n}}} \quad (23)$$

In this context, $\bar{d}$ represents the mean of the differences between paired observations (accuracy of models), s_d is the SD of these differences, and n is the number of observation pairs. The calculated t-statistics is then compared to a critical value from the t-distribution table, corresponding to the chosen significance level (e.g., 0.05), to decide whether to reject the null hypothesis. Model performance was statistically compared using the paired t-test, with the null hypothesis rejected when the p-value was less than 0.05, indicating a significant difference in predictive outcomes. In cases where the assumptions of the paired t-test (e.g., normality) were not satisfied, the Wilcoxon signed-rank test was employed as a non-parametric alternative. A p-value below 0.05 in the Wilcoxon test similarly indicated a statistically significant difference between models.

Following statistical testing, a comparative analysis of each model's strengths and limitations was conducted with respect to predict accuracy, interpretability, and robustness. Neural networks achieved the highest predictive accuracy, reflecting their ability to capture complex, non-linear relationships in endothelial responses under cyclic shear viscosity (CSV). However, their "black-box" nature limited interpretability, making it difficult to derive biologically actionable insights from the predictions. The depth and complexity of neural architectures enhanced performance but constrained transparency in linking features to endothelial mechanisms.

By contrast, RF offered a balanced trade-off between accuracy, interpretability, and robustness. Although their predictive accuracy was slightly lower than that of neural networks, RF demonstrated strong generalization to unseen data and provided valuable interpretability through feature importance analysis. With fuzzy logic integration, the RF-fuzzy model further enhanced robustness by capturing uncertainty and data variability through fuzzy membership functions. This approach enabled more nuanced identification of influential features, such as impedance modulus (|Z|), water contact angle (WCA), and electrochemical corrosion potential (Ecorr), which contributed most to endothelial adhesion strength. The combination of predictive reliability and interpretability positioned RF-fuzzy as a practical tool for vascular biology, particularly under complex biological conditions where uncertainty must be addressed. SVM also achieved high predictive performance, especially in handling high-dimensional datasets and distinguishing endothelial response classes through the RBF kernel. However, like neural networks, SVMs lacked direct feature importance measures, limiting their interpretability relative to RF. While accurate in modeling complex data structures, the inability to directly attribute predictions to specific variables constrained their biological applicability. To enhance interpretability across models, post-hoc techniques such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) were employed, particularly in RF-fuzzy and neural networks. These methods provided clearer insights into feature contributions, improving the biological relevance of model predictions. In summary, all three models neural networks, RF-fuzzy, and SVM performed effectively in predicting endothelial adhesion strength under CSV, but each excelled in different aspects. Neural networks provided superior accuracy, RF-fuzzy offered balanced performance with interpretability and robustness, and SVMs demonstrated strong predictive capacity in high-dimensional spaces but limited transparency. Importantly, RF-fuzzy contributed practical insights into endothelial mechanisms, supporting potential applications in optimizing drug delivery and therapeutic interventions in cardiovascular health.

Considering the senescence and CSV impact, several studies, such as those by (Costarelli *et al*. 2017, Wong *et al*. 2017) and (Guerrero *et al*. 2015), focus on senescence, a key factor in endothelial dysfunction, which could be accelerated under cyclic shear conditions. It's common to find that senescent cells exhibit changes in adhesion, migration, and gene expression, which may be reflected in the data of this study analyzed for adhesion strength under CSV. Changes in adhesion properties could provide insights

for model predictions regarding the influence of cyclic shear stress on endothelial function over time. For shear stress and endothelial behavior, the studies by (Mun *et al*. 2009) and (Laschober *et al*. 2010) highlight how ECs' behavior changes under different stress conditions, including laminar shear stress. The current study model similarly shows variations in endothelial adhesion, reflecting alterations in gene expression or mechanical properties as observed under static versus dynamic conditions. Based on the RF-fuzzy feature importance analysis, factors such as impedance modulus (|Z|), WCA, and Ecorr are frequently highlighted in studies as contributing significantly to cellular responses. These parameters are directly tied to the mechanical properties and surface characteristics of ECs under CSV, as seen in studies by (Shah *et al*. 2013) and (Rombouts *et al*. 2014), where exposure to varying conditions (e.g., radiation, age) modifies cellular behavior, potentially in ways similar to those under shear stress.

4. Conclusions

This research evaluated endothelial cell responses under cyclic shear viscosity (CSV) by comparing the performance of three machine learning models: Support Vector Machines, Random Forests with fuzzy integration, and neural networks. Model performance was assessed using multiple metrics, including accuracy, precision, recall, F1 score, (MAE), and (MSE), to ensure both predictive reliability and generalizability. Neural networks achieved the highest predictive accuracy (86.36%), reflecting their ability to model complex, non-linear patterns in the data; however, their limited interpretability constrained biological insight. The RF-fuzzy model achieved an accuracy of 85.24% and offered a balance between predictive capacity and interpretability. Feature importance analysis highlighted impedance modulus (| *Z* |), WCA, and electrochemical corrosion potential (Ecorr) as critical determinants of endothelial adhesion, with fuzzy integration further enhancing robustness against data variability. SVMs also performed effectively (82.50% accuracy), particularly in high-dimensional spaces, though their lack of direct feature importance limited biological interpretability. The findings underscore the importance of interpretability in applying machine learning to biological systems. While neural networks provided the highest accuracy, the RF-fuzzy approach offered the best trade-off between predictive accuracy and explanatory power, enabling actionable insights into endothelial behavior. This study demonstrates the value of integrating advanced ML techniques in vascular biology, contributing both methodological advancements and practical implications for optimizing drug delivery strategies and therapeutic interventions in cardiovascular health.

Acknowledgment

The authors acknowledge the Wenzhou Key Laboratory of Cardiopulmonary and Brain Resuscitation and Rehabilitation Application Transformation.

Funding

This work is supported by Wenzhou Major Scientific and Technological Innovation Project (ZY2024025).

References

Abood, E.A., M.H. Abdallah, M. Alsaadi, H. Imran, L.F.A. Bernardo, D. De Domenico and S.N. Henedy (2024), "Machine learning-based prediction models for punching shear strength of fiber-reinforced polymer reinforced concrete slabs using a gradient-boosted regression tree", *Materials*, **17**(16), 3964. https://doi.org/10.3390/ma17163964

AbuShanab, Y., W.A. Al-Ammari, S. Gowid and A.K. Sleiti (2023), "Accurate prediction of dynamic viscosity of polyalpha-olefin boron nitride nanofluids using machine learning", *Heliyon*, **9**(6). https://doi.org/10.1016/j.heliyon.2023.e16716

Amelio, I., F. Cutruzzolá, A. Antonov, M. Agostini and G. Melino (2014), "Serine and glycine metabolism in cancer", *Trends Biochem. Sci*., **39**(4), 191-198. https://doi.org/10.1016/j.tibs.2014.02.004

Anderson, R., G.D. Richardson and J.F. Passos (2018), "Mechanisms driving the ageing heart", *Experim. Gerontol*., **109**, 5-15. https://doi.org/10.1016/j.exger.2017.10.015

Andriani, G.A., V.P. Almeida, F. Faggioli, M. Mauro, W.L. Tsai, L. Santambrogio, A. Maslov, M. Gadina, J. Campisi and J. Vijg (2016), "Whole Chromosome Instability induces senescence and promotes SASP", *Sci. Rep*., **6**(1), 35218. https://doi.org/10.1038/srep35218

Arevalo-Rodriguez, I., D. Buitrago-Garcia, D. Simancas-Racines, P. Zambrano-Achig, R. Del Campo, A. Ciapponi, O. Sued, L. Martinez-Garcia, A.W. Rutjes and N. Low (2020), "False-negative results of initial RT-PCR assays for COVID-19: A systematic review", *PloS One*, **15**(12), e0242958. https://doi.org/10.1371/journal.pone.0242958

Arevalo-Rodriguez, I., P. Seron, D. Buitrago-García, A. Ciapponi, A. Muriel, P. Zambrano-Achig, R. Del Campo, J.C. Galán-Montemayor, D. Simancas-Racines and J.A. Perez-Molina (2021), "Recommendations for SARS-CoV-2/COVID-19 testing: A scoping review of current guidance", *BMJ Open*. **11**(1), e043004. https://doi.org/10.1136/bmjopen-2020-043004

Armaghani, D.J., F. Mirzaei, M. Shariati, N.T. Trung, M. Shariati and D. Trnavac (2020), "Hybrid ANN-based techniques in predicting cohesion of sandy-soil combined with fiber", *Geomech. Eng*., **20**(3), 191-205. https://doi.org/10.12989/gae.2020.20.3.191

Avila, R.M., G. Camacho-Leon, M. Faytong-Haro, R.L. Merino-Alado, J.P. Carrillo, H. Mautong, D.A. Simancas-Racines and I. Cherrez-Ojeda (2024), "Case report: Primary cutaneous histoplasmosis in an immunocompetent patient after cosmetic injection of platelet-rich plasma treated with trimethoprim-sulfamethoxazole", *Am. J. Case Rep*., **25**, e942660-942661. https://doi.org/10.12659/AJCR.942660

Aviv, H., M.Y. Khan, J. Skurnick, K. Okuda, M. Kimura, J. Gardner, L. Priolo and A. Aviv (2001), "Age dependent aneuploidy and telomere length of the human vascular endothelium", *Atherosclerosis*, **159**(2), 281-287. https://doi.org/10.1016/S0021-9150(01)00506-8

Bao, M., G. Li, X. Huang, L. Tang, L. Dong and J. Li (2018), "Long noncoding RNA LINC00657 acting as a miR-590-3p sponge to facilitate low concentration oxidized low-density lipoprotein–induced angiogenesis", *Mol. Pharmacol*., **93**(4), 368-375. https://doi.org/10.1124/mol.117.110650

Barkhordari, M.S. and C. Qi (2025), "Prediction of zinc, cadmium, and arsenic in european soils using multi-end machine learning models", *J. Hazard. Mater*., **490**, 137800.

https://doi.org/10.1016/j.jhazmat.2025.137800
Bonfill, X., D. Osorio, M. Posso, I. Solà, G. Rada, A. Torres, M. García Dieguez, M. Piña-Pozas, L. Díaz-García and M. Tristán (2015), "Identification of biomedical journals in Spain and Latin America", *Health Inform. Libr. J.*, **32**(4), 276-286. https://doi.org/10.1111/hir.12110
Breitling, R., P. Armengaud, A. Amtmann and P. Herzyk (2004), "Rank products: A simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments", *FEBS Lett.*, **573**(1-3), 83-92. https://doi.org/10.1016/j.febslet.2004.07.055
Cadena-Ullauri, S., P. Guevara-Ramírez, V.A. Ruiz-Pozo, R. Tamayo-Trujillo, E. Paz-Cruz, D. Simancas-Racines, R. Ibarra-Castillo, J.L. Laso-Bayas and A.K. Zambrano (2024), "Genomic analysis of an Ecuadorian individual carrying an SCN5A rare variant", *BMC Cardiovasc. Disord.*, **24**(1), 388. https://doi.org/10.1186/s12872-024-04049-w
Cadena-Ullauri, S., P. Guevara-Ramírez, V.A. Ruiz-Pozo, R. Tamayo-Trujillo, E. Paz-Cruz, R. Zambrano-Villacres, D. Simancas-Racines and A.K. Zambrano (2024), "The effect of intermittent fasting on microbiota as a therapeutic approach in obesity", *Front. Nutr.*, **11**, 1393292. https://doi.org/10.3389/fnut.2024.1393292
Calderón, J., A. Cherrez, G.D. Ramón, O.L. Jove, A. Baptist, E. Matos, B.M. Maciel, E. Calero, M. Sanchez-Borges and S. Cherrez (2017), "Information and communication technology use in asthmatic patients: A cross-sectional study in Latin America", *ERJ Open Res.*, **3**(3). https://doi.org/10.1183/23120541.00005-2017
Canbek, G. (2023), "BenchMetrics Prob: benchmarking of probabilistic error/loss performance evaluation instruments for binary classification problems", *Int. J. Mach. Learn. Cybernet.*, **14**(9), 3161-3191. https://doi.org/10.1007/s13042-023-01826-5
Chapela, S.P., A. Simancas-Racines, F. Ceriani, A.L.N. Martinuzzi, M.P. Russo, A.K. Zambrano, D. Simancas-Racines, L. Verde, G. Muscogiuri and C.S. Katsanos (2024), "Obesity and Obesity-Related Thyroid Dysfunction: Any Potential Role for the Very Low-Calorie Ketogenic Diet (VLCKD)?", *Curr. Nutr. Rep.*, **13**(2), 194-213. https://doi.org/10.1007/s13668-024-00528-w
Chapela, S.P., D. Simancas-Racines, M. Montalvan, E. Frias-Toral, A. Simancas-Racines, G. Muscogiuri, L. Barrea, G. Sarno, P.I. Martínez and M.J. Reberendo (2023), "Signals for muscular protein turnover and insulin resistance in critically ill patients: A narrative review", *Nutrients*, **15**(5), 1071. https://doi.org/10.3390/nu15051071
Chen, S., K. Zhou, H. Shang, M. Du, L. Wu and Y. Chen (2023), "Effects of concurrent aerobic and resistance training on vascular health in type 2 diabetes: a systematic review and meta-analysis", *Front. Endocrinol.*, **14**, 1216962. https://doi.org/10.3389/fendo.2023.1216962
Chen, Z., Y. Wen, W. Li, J. Bai, P. Zhou, Q. He and Z. Deng (2025), "Geriatric nutritional risk index as a predictor of mortality in women with chronic inflammatory airway disease: evidence from NHANES 1999–2018", *Front. Nutr.*, **12**, 1547952. https://doi.org/10.3389/fnut.2025.1547952
Chérrez-Ojeda, I., J.C. Calderón, A. Fernández García, D.B. Jeffe, I. Santoro, E. Vanegas, A. Cherrez, J. Cano, F. Betancourt and D. Simancas-Racines (2018), "Obstructive sleep apnea knowledge and attitudes among recent medical graduates training in Ecuador", *Multidiscipl. Respir. Med.*, **13**, 1-8. https://doi.org/10.1186/s40248-018-0117-8
Chérrez-Ojeda, I., M. Felix, V.L. Mata, E. Vanegas, D. Simancas-Racines, M. Aguilar, A.W. Gavilanes, P. Chedraui and C. Vera (2020), "Use and perceptions of information and communication technologies among Ecuadorian nurses: a cross-sectional study", *Open Nurs. J.*, **14**(1). http://doi.org/10.2174/1874434602014010008
Cherrez-Ojeda, I., E. Vanegas, M. Felix, V.L. Mata, F.M. Jiménez, M. Sanchez, D. Simancas-Racines, S. Cherrez, A.W. Gavilanes and J. Eschrich (2020), "Frequency of use, perceptions and barriers of information and communication technologies among Latin American physicians: An Ecuadorian cross-sectional study", *J. Multidiscipl. Healthcare*, 259-269. https://doi.org/10.2147/JMDH.S246253
Childs, B.G., H. Li and J.M. Van Deursen (2018), "Senescent cells: A therapeutic target for cardiovascular disease", *J. Clinic. Investigat.*, **128**(4), 1217-1228. https://doi.org/10.1172/JCI95146.
Choi, W., T. Choi and S. Heo (2023), "A comparative study of automated machine learning platforms for exercise anthropometry-based typology analysis: Performance evaluation of AWS SageMaker, GCP VertexAI, and MS Azure", *Bioengineering*, **10**(8), 891. https://doi.org/10.3390/bioengineering10080891
Clarke, R., H.W. Ressom, A. Wang, J. Xuan, M.C. Liu, E.A. Gehan and Y. Wang (2008), "The properties of high-dimensional data spaces: implications for exploring gene and protein expression data", *Nature Rev. Cancer*, **8**(1), 37-49. https://doi.org/10.1038/nrc2294
Coppé, J.P., P.Y. Desprez, A. Krtolica and J. Campisi (2010), "The senescence-associated secretory phenotype: the dark side of tumor suppression", *Annual Rev. Pathol. Mech. Disease*, **5**(1), 99-118. https://doi.org/10.1146/annurev-pathol-121808-102144
Cortes, M.B., R.S.N. da Silva, P.C. de Oliveira, D.S. da Silva, M.C.C. Irigoyen, G. Waclawovsky and M.I. Schaun (2023), "Effect of aerobic and resistance exercise training on endothelial function in individuals with overweight and obesity: A systematic review with meta-analysis of randomized clinical trials", *Sci. Rep.*, **13**(1), 11826. https://doi.org/10.1038/s41598-023-38603-x
Costarelli, L., R. Giacconi, M. Malavolta, A. Basso, F. Piacenza, M. Provinciali, M.G. Maggio, A. Corsonello and F. Lattanzio (2017), "Different transcriptional profiling between senescent and non-senescent human coronary artery endothelial cells (HCAECs) by Omeprazole and Lansoprazole treatment", *Biogerontology*, **18**, 217-236. https://doi.org/10.1007/s10522-016-9675-3
Daviran, M., A. Maghsoudi, R. Ghezelbash and B. Pradhan (2021), "A new strategy for spatial predictive mapping of mineral prospectivity: Automated hyperparameter tuning of random forest approach", *Comput., Geosci.*, **148,** 104688. https://doi.org/10.1016/j.cageo.2021.104688
Du, H., J. Sarno and H.S. Taylor (2010), "HOXA10 inhibits Kruppel-like factor 9 expression in the human endometrial epithelium", *Biol. Reprod.*, **83**(2), 205-211. https://doi.org/10.1095/biolreprod.110.083980
Flammer, A.J., T. Anderson, D.S. Celermajer, M.A. Creager, J. Deanfield, P. Ganz, N.M. Hamburg, T.F. Lüscher, M. Shechter and S. Taddei (2012), "The assessment of endothelial function: From research into clinical practice", *Circulation*, **126**(6), 753-767. https://doi.org/10.1161/CIRCULATIONAHA.112.093245
Franco, J.V.A., M. Arancibia, D. Simancas-Racines and E. Madrid (2018), "Síntesis de información biomédica: revisiones narrativas, revisiones sistemáticas y estructuras emergentes", *Medwave*, **18**(7). https://doi.org/10.5867/medwave.2018.07.7354
Franco, J.V.A., M. Arancibia, D. Simancas-Racines and E. Madrid (2018), "Syntheses of biomedical information: narrative reviews, systematic reviews and emerging formats", *Medwave*, **18**(7). https://doi.org/10.5867/medwave.2018.07.7354
Fuentes, R., D. Osorio, J.E. Hernandez, D. Simancas-Racines, M.J. Martinez-Zapata and X.B. Cosp (2018), "Surgery versus stereotactic radiotherapy for people with single or solitary brain metastasis", *Cochrane Database System. Rev.*, (8). https://doi.org/10.1002/14651858.CD012086.pub2

Ghezelbash, R., A. Maghsoudi and M. Daviran (2019), "Combination of multifractal geostatistical interpolation and spectrum–area (S–A) fractal model for Cu–Au geochemical prospects in Feizabad district, NE Iran", *Arab. J. Geosci.*, **12**, 1-14. https://doi.org/10.1007/s12517-019-4318-z

Godo, S. and H. Shimokawa (2017), "Endothelial functions", Arteriosclerosis, thrombosis, and vascular biology **37**(9), e108-e114. https://doi.org/10.1161/ATVBAHA.117.309813

Green, D.J., H. Jones, D. Thijssen, N. Cable and G. Atkinson (2011), "Flow-mediated dilation and cardiovascular event prediction: does nitric oxide matter?", *Hypertension*, **57**(3), 363-369. https://doi.org/10.1161/HYPERTENSIONAHA.110.167015

Guerrero, A., C. Iglesias, S. Raguz, E. Floridia, J. Gil, C. M. Pombo and J. Zalvide (2015), "The cerebral cavernous malformation 3 gene is necessary for senescence induction", *Aging Cell*, **14**(2), 274-283. https://doi.org/10.1111/acel.12316

Guevara-Ramírez, P., S. Cadena-Ullauri, V.A. Ruiz-Pozo, R. Tamayo-Trujillo, E. Paz-Cruz, D. Simancas-Racines and A.K. Zambrano (2022), "Genetics, genomics, and diet interactions in obesity in the Latin American environment", *Front. Nutr.*, **9**, 1063286. https://doi.org/10.3389/fnut.2022.1063286

Guevara-Ramírez, P., E. Paz-Cruz, S. Cadena-Ullauri, V.A. Ruiz-Pozo, R. Tamayo-Trujillo, M.L. Felix, D. Simancas-Racines and A.K. Zambrano (2023), "Molecular pathways and nutrigenomic review of insulin resistance development in gestational diabetes mellitus", *Front. Nutr.*, **10**, 1228703. https://doi.org/10.3389/fnut.2023.1228703

Gytz, H., M.F. Hansen, S. Skovbjerg, A.C. Kristensen, S. Hørlyck, M.B. Jensen, M. Fredborg, L.D. Markert, N.A. McMillan and E.I. Christensen (2017), "Apoptotic properties of the type 1 interferon induced family of human mitochondrial membrane ISG12 proteins", *Biol. Cell.*, **109**(2), 94-112. https://doi.org/10.1111/boc.201600034

Hastie, T. and W. Stuetzle (1989), "Principal curves", J. Am. Statist. Assoc., **84**(406), 502-516. https://doi.org/10.1080/01621459.1989.10478797

Hebbel, R.P., P. Wei, L. Milbauer, M.T. Corban, A. Solovey, J. Kiley, J. Pattee, L.O. Lerman, W. Pan and A. Lerman (2020), "Abnormal endothelial gene expression associated with early coronary atherosclerosis", *J. Am. Heart Assoc.*, **9**(14), e016134. https://doi.org/10.1161/JAHA.120.016134

Hidalgo, R., A.J. Martí-Carvajal, J.S. Kwong, D. Simancas-Racines and S. Nicola (2012), "Pharmacological interventions for treating heart failure in patients with chagas cardiomyopathy", *Cochrane Database Syst. Rev.*, (11). https://doi.org/10.1002/14651858.CD009077.pub3

Hughey, J.J. and A.J. Butte (2015), "Robust meta-analysis of gene expression using the elastic net", *Nucleic Acids Res.*, **43**(12), e79-e79. https://doi.org/10.1093/nar/gkv229

Inaba, Y., J.A. Chen and S.R. Bergmann (2010), "Prediction of future cardiovascular outcomes by flow-mediated vasodilatation of brachial artery: A meta-analysis", *Int. J. Cardiovasc. Imag.*, **26**, 631-640. https://doi.org/10.1007/s10554-010-9616-1

Jaffe, E.A., R.L. Nachman, C.G. Becker and C.R. Minick (1973), "Culture of human endothelial cells derived from umbilical veins. Identification by morphologic and immunologic criteria", *J. Clinic. Invest.*, **52**(11), 2745-2756. https://doi.org/10.1172/JCI107470

Jahandari, S., Z. Tao, M. Saberian, M. Shariati, J. Li, M. Abolhasani, M. Kazemi, A. Rahmani and M. Rashidi (2022), "Geotechnical properties of lime-geogrid improved clayey subgrade under various moisture conditions", *Road Mater. Pave. Des.*, **23**(9), 2057-2075. https://doi.org/10.1080/14680629.2021.1950816

Jia, G., A.R. Aroor, C. Jia and J.R. Sowers (2019), "Endothelial cell senescence in aging-related vascular dysfunction", *Biochimica et Biophysica Acta (BBA)*, **1865**(7), 1802-1809. https://doi.org/10.1016/j.bbadis.2018.08.008

Jia, Q., H. Huang, Z. Tong, L. Fang, Q. Jia, S. Zhu, Y. Zheng, S. Guan, D. Bian and H. Yu (2024), "Endothelium-Mimicking with NO-generating coating on bioabsorbable magnesium alloy for improving corrosion resistance and biological responses of vascular stents", *Chem. Eng. J.*, **490**, 151344. https://doi.org/10.1016/j.cej.2024.151344

Johnson, W.E., C. Li and A. Rabinovic (2007), "Adjusting batch effects in microarray expression data using empirical Bayes methods", *Biostatistics*, **8**(1), 118-127. https://doi.org/10.1093/biostatistics/kxj037

Jong, H.L., M.R. Mustafa, P.M. Vanhoutte, S. AbuBakar and P.F. Wong (2013), "MicroRNA 299-3p modulates replicative senescence in endothelial cells", *Physiol. Genom.*, **45**(7), 256-267. https://doi.org/10.1152/physiolgenomics.00071.2012

Kasahun, M. and A. Legesse (2024), "Machine learning for urban land use/ cover mapping: Comparison of artificial neural network, random forest and support vector machine, a case study of Dilla town", *Heliyon*, **10**(20), e39146. https://doi.org/10.1016/j.heliyon.2024.e39146

Katebi, J., M. Shoaei-parchin, M. Shariati, N.T. Trung and M. Khorami (2020), "Developed comparative analysis of metaheuristic optimization algorithms for optimal active control of structures", *Eng. Comput.*, **36**, 1539-1558. https://doi.org/10.1007/s00366-019-00780-7

Kernbach, J.M. and V.E. Staartjes (2022), "Foundations of machine learning-based clinical prediction modeling: Part II—Generalization and overfitting", *Mach. Learn. Clinic. Neurosci.*, 15-21. https://doi.org/10.1007/978-3-030-85292-4_3

Kim, J.H. (2024), "A study on the water content in distribution pole transformer using random forest model", *Comput. Electr. Eng.*, **120**, 109823. https://doi.org/10.1016/j.compeleceng.2024.109823

Laschober, G.T., D. Ruli, E. Hofer, C. Muck, D. Carmona-Gutierrez, J. Ring, E. Hutter, C. Ruckenstuhl, L. Micutkova and R. Brunauer (2010), "Identification of evolutionarily conserved genetic regulators of cellular aging", *Aging Cell*, **9**(6), 1084-1097. https://doi.org/10.1111/j.1474-9726.2010.00637.x

Lew, L.A., T.S. Ethier and K.E. Pyke (2022), "The impact of exercise training on endothelial function in postmenopausal women: A systematic review", *Experim. Physiol.*, **107**(12), 1388-1421. https://doi.org/10.1113/EP090702

Li, D., A. Toghroli, M. Shariati, F. Sajedi, D.T. Bui, P. Kianmehr, E.T. Mohamad and M. Khorami (2019), "Application of polymer, silica-fume and crushed rubber in the production of Pervious concrete", *Smart Struct. Syst.*, **23**(2), 207-214. https://doi.org/10.12989/sss.2019.23.2.207

Li, H., Z. Wang, Z. Guan, J. Miao, W. Li, P. Yu and C. Molina Jimenez (2024), "UCFNNet: Ulcerative colitis evaluation based on fine-grained lesion learner and noise suppression gating", *Comput. Meth. Prog. Biomed.*, **247**, 108080. https://doi.org/10.1016/j.cmpb.2024.108080

Liu, Z., L. Si, S. Shi, J. Li, J. Zhu, W. H. Lee and G. Wang (2024), "Classification of three anesthesia stages based on near-infrared spectroscopy signals", *IEEE J. Biomed. Health Inform.*, **28**(9), 5270-5279. https://doi.org/10.1109/JBHI.2024.3409163

López, M., M.J. Paucar, M. Vega, D. Izurieta, H. Acosta, D. Simancas-Racines and J. Angamarca (2023), "Factores asociados al embarazo adolescente en Ecuador y el rol del primer nivel de atención en salud: Un análisis a partir del instrumento ENSANUT 2018", *Práctica Familiar Rural*, **8**(3).

Marti-Carvajal, A.J., D. Simancas-Racines, V. Anand and S.I. Bangdiwala (2015), "Prophylactic lidocaine for myocardial infarction", *Cochrane Database Syst. Rev.*, (8). https://doi.org/10.1002/14651858.CD008553.pub2

Martí-Carvajal, A.J., C. Gluud, S. Nicola, D. Simancas-Racines, L.

Reveiz, P. Oliva and J. Cedeño-Taborda (2015), "Growth factors for treating diabetic foot ulcers", *Cochrane Database Syst. Rev*., (10). https://doi.org/10.1002/14651858.CD008548.pub2

Martí-Carvajal, A.J., D. Simancas-Racines and B.S. Peña-González (2015), "Prolonged storage of packed red blood cells for blood transfusion", *Cochrane Database Syst. Rev*., (7). https://doi.org/10.1002/14651858.CD009330

Mayorga-Ramos, A., C. Barba-Ostria, D. Simancas-Racines and L. P. Guamán (2022), "Protective role of butyrate in obesity and diabetes: New insights", *Front. Nutr*., **9**, 1067647. https://doi.org/10.3389/fnut.2022.1067647

Mohammadhassani, M., H. Nezamabadi-Pour, M. Suhatril and M. Shariati (2014), "An evolutionary fuzzy modelling approach and comparison of different methods for shear strength prediction of high-strength concrete beams without stirrups", *Smart. Struct. Syst*., **14**(5), 785-809. http://doi.org/10.12989/sss.2014.14.5.785

Mohammadhassani, M., M. Suhatril, M. Shariati and F. Ghanbari (2013), "Ductility and strength assessment of HSC beams with varying of tensile reinforcement ratios", *Struct. Eng. Mech*., **48**(6), 833-848. https://doi.org/10.12989/sem.2013.48.6.833

Molano-Franco, D., I. Arevalo-Rodriguez, A. Muriel, L. del Campo-Albendea, S. Fernández-García, A. Alvarez-Méndez, D. Simancas-Racines, A. Viteri, G. Sanchez and B. Fernandez-Felix (2023), "Basal procalcitonin, C-reactive protein, interleukin-6, and presepsin for prediction of mortality in critically ill septic patients: A systematic review and meta-analysis", *Diagnos. Prognost. Res*., **7**(1), 15. https://doi.org/10.1186/s41512-023-00152-2

Montero-Oleas, N., I. Arevalo-Rodriguez, S. Nuñez-González, A. Viteri-García and D. Simancas-Racines (2020), "Therapeutic use of cannabis and cannabinoids: An evidence mapping and appraisal of systematic reviews", *BMC Complement. Med. Therap*., **20**, 1-15. https://doi.org/10.1186/s12906-019-2803-2

Montero-Oleas, N., S. Nunez-Gonzalez and D. Simancas-Racines (2017), "The remarkable geographical pattern of gastric cancer mortality in Ecuador", *Cancer Epidemiol*., **51**, 92-97. https://doi.org/10.1016/j.canep.2017.10.014

Montesinos-Guevara, C., D. Buitrago-Garcia, M.L. Felix, C.V. Guerra, R. Hidalgo, M.J. Martinez-Zapata and D. Simancas-Racines (2022), "Vaccines for the common cold", *Cochrane Database Syst. Rev*., (12). https://doi.org/10.1002/14651858.CD002190.pub5

Montesinos-Guevara, C. and D. Simancas-Racines (2019), "Reflexiones sobre bioética y la atención primaria de la salud en el Ecuador", *Práctica Familiar Rural*, **4**(3).

Morales, C.G., K.S. Arévalo, V.G. Cevallos, D.S. Racines and C.M. Guevara (2021), "Reapertura de escuelas durante la pandemia de COVID-19", *Práctica Familiar Rural*, **6**(1), 6.

Mun, G.I., S.J. Lee, S.M. An, I.K. Kim and Y.C. Boo (2009), "Differential gene expression in young and senescent endothelial cells under static and laminar shear stress conditions", *Free Radical Biol. Med*., **47**(3), 291-299. https://doi.org/10.1016/j.freeradbiomed.2009.04.032

Muscogiuri, G., L. Verde, E. Frias-Toral, C. Reytor-González, G. Annunziata, M. Proganò, S. Savastano, D. Simancas-Racines, A. Colao and L. Barrea (2024), "Weight loss, changes in body composition and inflammatory status after a very low-energy ketogenic therapy (VLEKT), does gender matter?", *J. Translat. Med*., **22**(1), 949. https://doi.org/10.1186/s12967-024-05733-3

Naghipour, M., G. Yousofizinsaz and M. Shariati (2020), "Experimental study on axial compressive behavior of welded built-up CFT stub columns made by cold-formed sections with different welding lines", *Steel Compos. Struct*, **34**(3), 347-359. https://doi.org/10.12989/scs.2020.34.3.347

Nugroho, A., E.M. Yuniarno and M.H. Purnomo (2023), "ARKOMA dataset: An open-source dataset to develop neural networks-based inverse kinematics model for NAO robot arms", *Data in Brief*, **51**, 109727. https://doi.org/10.1016/j.dib.2023.109727

Núñez-González, S., E. Bedoya, D. Simancas-Racines and C. Gault (2020), "Spatial clusters and temporal trends of malignant melanoma mortality in Ecuador", *SAGE Open Med*., **8**, 2050312120918285. https://doi.org/10.1177/2050312120918285

Núñez-González, S., A. Delgado-Ron and D. Simancas-Racines (2020), "Tendencias y análisis espacio-temporal de la mortalidad por diabetes mellitus en Ecuador, 2001-2016", *Revista Cubana de Salud Pública*, **46**, e1314.

Núñez-González, S., J.A. Delgado-Ron, C. Gault, A. Lara-Vinueza, D. Calle-Celi, R. Porreca and D. Simancas-Racines (2020), "Overview of "systematic reviews" of the built environment's effects on mental health", *J. Environ. Publ. Health*, **2020**(1), 9523127. https://doi.org/10.1155/2020/9523127

Núñez-González, S., C. Gault, P. Granja and D. Simancas-Racines (2020), "Spatial patterns of leptospirosis in Ecuador, 2013–2018", *T Royal Soc. Tropic. Med. Hygiene*, **114**(7), 545-548. https://doi.org/10.1093/trstmh/trz124

Park, H.S. and S.Y. Kim (2021), "Endothelial cell senescence: A machine learning-based meta-analysis of transcriptomic studies", *Age. Res. Rev*., **65**, 101213. https://doi.org/10.1016/j.arr.2020.101213

Parvandeh, S., H.W. Yeh, M.P. Paulus and B.A. McKinney (2020), "Consensus features nested cross-validation", *Bioinformatics* **36**(10), 3093-3098. https://doi.org/10.1093/bioinformatics/btaa046

Paz-Cruz, E., S. Cadena-Ullauri, P. Guevara-Ramírez, V.A. Ruiz-Pozo, R. Tamayo-Trujillo, D. Simancas-Racines and A.K. Zambrano (2023), "Thyroid cancer in Ecuador: A genetic variants review and a cross-sectional population-based analysis before and after COVID-19 pandemic", *Heliyon*, **10**(1). https://doi.org/10.1016/j.heliyon.2023.e23964

Paz-Cruz, E., S. Cadena-Ullauri, P. Guevara-Ramírez, V.A. Ruiz-Pozo, R. Tamayo-Trujillo, D. Simancas-Racines and A.K. Zambrano (2024), "Response to a letter to the editor,"Studying the incidence of thyroid cancer in Ecuador: 2016–2021", *Heliyon*, **10**(10). https://doi.org/10.1016/j.heliyon.2024.e30963

Pearson, M. and N. Smart (2017), "Effect of exercise training on endothelial function in heart failure patients: A systematic review meta-analysis", *Int. J. Cardiol*., **231**, 234-243. https://doi.org/10.1016/j.ijcard.2016.12.145

Peng, L., W. Bo, H. Yang and X. Li (2025), "Deep learning-based image compression for enhanced hyperspectral processing in the protection of stone cultural relics", *Exp. Syst. Appl*., **271**, 126691. https://doi.org/10.1016/j.eswa.2025.126691

Qiu, S., X. Cai, H. Yin, Z. Sun, M. Zügel, J. M. Steinacker and U. Schumann (2018), "Exercise training and endothelial function in patients with type 2 diabetes: A meta-analysis", *Cardiovasc. Diabetol*., **17**, 1-12. https://doi.org/10.1186/s12933-018-0711-2

Ras, R.T., M.T. Streppel, R. Draijer and P.L. Zock (2013), "Flow-mediated dilation and cardiovascular risk prediction: A systematic review with meta-analysis", *Int. J. Cardiol*., **168**(1), 344-351. https://doi.org/10.1016/j.ijcard.2012.09.047

Reytor-González, C., J.M. Parise-Vasco, N. González, A. Simancas-Racines, R. Zambrano-Villacres, A. K. Zambrano and D. Simancas-Racines (2024), "Obesity and periodontitis: A comprehensive review of their interconnected pathophysiology and clinical implications", *Front. Nutr*., **11**, 1440216. https://doi.org/10.3389/fnut.2024.1440216

Robeson, S.M. and C.J. Willmott (2023), "Decomposition of the mean absolute error (MAE) into systematic and unsystematic components", *PloS One*, **18**(2), e0279774. https://doi.org/10.1371/journal.pone.0279774

Rombouts, C., A. Aerts, R. Quintens, B. Baselet, H. El-Saghire, M. Harms-Ringdahl, S. Haghdoost, A. Janssen, A. Michaux and R.

Yentrapalli (2014), "Transcriptomic profiling suggests a role for IGFBP5 in premature senescence of endothelial cells after chronic low dose rate irradiation", *Int. J. Radiat. Biol.*, https://doi.org/10.3109/09553002.2014.905724

Rubio-Gallegos, F., S. Núñez-González, C. Gault, D. Simancas-Racines and E. Basantes-García (2019), "McGregor inguinal flap for coverage of large soft tissue losses due to high-voltage electrical burns in the upper limb: A retrospective study", *Int. J. Burns Trauma*, **9**(3), 52.

Ruiz-Pozo, V.A., P. Guevara-Ramírez, E. Paz-Cruz, R. Tamayo-Trujillo, S. Cadena-Ullauri, E. Frias-Toral, D. Simancas-Racines, Y. Altuna-Roshkova, C. Reytor-González and A.K. Zambrano (2024), "The role of the Mediterranean diet in prediabetes management and prevention: A review of molecular mechanisms and clinical outcomes", *Food Agric. Immunol.*, **35**(1), 2398042. https://doi.org/10.1080/09540105.2024.2398042

Ruiz-Pozo, V.A., R. Tamayo-Trujillo, S. Cadena-Ullauri, E. Frias-Toral, P. Guevara-Ramírez, E. Paz-Cruz, S. Chapela, M. Montalván, T. Morales-López and D. Simancas-Racines (2023), "The molecular mechanisms of the relationship between insulin resistance and Parkinson's disease pathogenesis", *Nutrients*, **15**(16), 3585. https://doi.org/10.3390/nu15163585

Safa, M., P.A. Sari, M. Shariati, M. Suhatril, N.T. Trung, K. Wakil and M. Khorami (2020), "Development of neuro-fuzzy and neuro-bee predictive models for prediction of the safety factor of eco-protection slopes", *Physica A*, **550**, 124046. https://doi.org/10.1016/j.physa.2019.124046

Salazar, J., J. Pérez-Bracchiglione, K. Salas-Gama, A. Antequera, A. Auladell-Rispau, R. Dorantes-Romandía, A.G. Meade, M.J. Quintana, C. Requeijo and G. Rodríguez-Grijalva (2021), "Efficacy of systemic oncological treatments in patients with advanced pancreatic cancer at high risk of dying in the short or medium-term: Overview of systematic reviews", *Eur. J. Cancer*, **154**, 82-91. https://doi.org/10.1016/j.ejca.2021.05.034

Sanlı, T., Ç. Sıcakyüz and O. Yüregir (2020), "Comparison of the accuracy of classification algorithms on three data-sets in data mining: Example of 20 classes", *Int. J. Eng. Sci. Technol.*, **12**(3), 81-89. https://doi.org/10.4314/ijest.v12i3.8

Sedghi, Y., Y. Zandi, M. Shariati, E. Ahmadi, V.M. Azar, A. Toghroli, M. Safa, E.T. Mohamad, M. Khorami and K. Wakil (2018), "Application of ANFIS technique on performance of C and L shaped angle shear connectors", *Smart Struct. Syst.*, **22**(3), 335-340. https://doi.org/10.12989/sss.2018.22.3.335

Shah, P.P., G. Donahue, G.L. Otte, B.C. Capell, D.M. Nelson, K. Cao, V. Aggarwala, H.A. Cruickshanks, T.S. Rai and T. McBryan (2013), "Lamin B1 depletion in senescent cells triggers large-scale changes in gene expression and the chromatin landscape", *Genes Develop.*, **27**(16), 1787-1799. https://doi,org/10.1101/gad.223834.113

Shao, Y.H., N.Y. Deng and Z.M. Yang (2012), "Least squares recursive projection twin support vector machine for classification", *Pattern Recog.*, **45**(6), 2299-2307. https://doi.org/10.1016/j.patcog.2011.11.028

Shariati, M., S.M. Davoodnabi, A. Toghroli, Z. Kong and A. Shariati (2021), "Hybridization of metaheuristic algorithms with adaptive neuro-fuzzy inference system to predict load-slip behavior of angle shear connectors at elevated temperatures", *Compos. Struct.*, **278**, 114524. https://doi.org/10.1016/j.compstruct.2021.114524

Shariati, M., M.S. Mafipour, B. Ghahremani, F. Azarhomayun, M. Ahmadi, N.T. Trung and A. Shariati (2022), "A novel hybrid extreme learning machine–grey wolf optimizer (ELM-GWO) model to predict compressive strength of concrete with partial replacements for cement", *Eng. Comput.*, 1-23. https://doi.org/10.1007/s00366-020-01081-0

Shariati, M., M.S. Mafipour, J.H. Haido, S.T. Yousif, A. Toghroli, N.T. Trung and A. Shariati (2020), "Identification of the most influencing parameters on the properties of corroded concrete beams using an Adaptive Neuro-Fuzzy Inference System (ANFIS)", *Steel Compos Struct.*, **34**(1), 155. https://doi.org/10.12989/scs.2020.34.1.155

Shariati, M., M.S. Mafipour, P. Mehrabi, A. Shariati, A. Toghroli, N.T. Trung and M.N. Salih (2021), "A novel approach to predict shear strength of tilted angle connectors using artificial intelligence techniques", *Eng. Comput.*, **37**, 2089-2109. https://doi.org/10.1007/s00366-019-00930-x

Shariati, M., F. Tahmasbi, P. Mehrabi, A. Bahadori and A. Toghroli (2020), "Monotonic behavior of C and L shaped angle shear connectors within steel-concrete composite beams: An experimental investigation", *Steel Compos Struct.*, **35**(2), 237-247. https://doi.org/10.12989/scs.2020.35.2.237

Shen, C., M.Y. Tsai, L. Chen, S. Li, D. Nguyen, J. Wang, S.B. Jiang and X. Jia (2020), "On the robustness of deep learning-based lung-nodule classification for CT images with respect to image noise", *Phys. Med. Biol.*, **65**(24), 245037. https://doi.org/10.1088/1361-6560/abc812

Shivgulam, M.E., H. Liu, B.D. Schwartz, J.E. Langley, N.W. Bray, D.S. Kimmerly and M.W. O'Brien (2023), "Impact of exercise training interventions on flow-mediated dilation in adults: An umbrella review", *Sports Med.*, **53**(6), 1161-1174. https://doi.org/10.1007/s40279-023-01837-w

Simancas-Racines, A., S. Cadena-Ullauri, P. Guevara-Ramírez, A. K. Zambrano and D. Simancas-Racines (2023), "Avian influenza: Strategies to manage an outbreak", *Pathogens*, **12**(4), 610. https://doi.org/10.3390/pathogens12040610

Simancas-Racines, D., I. Arevalo-Rodriguez, G. Urrutia, D. Buitrago-Garcia, S. Núñez-González, M. J. Martínez-Zapata, E. Madrid, X. Bonfill and R. Hidalgo-Ottolenghi (2019), "Leukodepleted packed red blood cells transfusion in patients undergoing major cardiovascular surgical procedure: Systematic review and meta-analysis", *Cardiol. Res. Pract.*, **2019**(1), 7543917. https://doi.org/10.1155/2019/7543917

Simancas-Racines, D., I. Arevalo-Rodriguez, D. Osorio, J.V. Franco, Y. Xu and R. Hidalgo (2018), "Interventions for treating acute high altitude illness", *Cochrane Database Syst. Rev.*, (6). https://doi.org/10.1002/14651858.CD009567.pub2

Simancas-Racines, D., J.V. Franco, C.V. Guerra, M.L. Felix, R. Hidalgo and M.J. Martinez-Zapata (2017), "Vaccines for the common cold", *Cochrane Database Syst. Rev.*, (5). https://doi.org/10.1002/14651858.CD002190.pub5

Simancas-Racines, D., N. Montero-Oleas, R.W. Vernooij, I. Arevalo-Rodriguez, P. Fuentes, I. Gich, R. Hidalgo, M.J. Martinez-Zapata, X. Bonfill and P. Alonso-Coello (2019), "Quality of clinical practice guidelines about red blood cell transfusion", *J. Evidence Based Med.*, **12**(2), 113-124. https://doi.org/10.1111/jebm.12330

Simancas-Racines, D., D. Osorio, A.J. Martí-Carvajal and I. Arevalo-Rodriguez (2015), "Leukoreduction for the prevention of adverse reactions from allogeneic blood transfusion", *Cochrane Database Syst. Rev.*, (12). https://doi.org/10.1002/14651858.CD009745.pub2

Song, K., H. Yang, L. Chen and M. Jaboyedoff (2024), "Step-like displacement prediction and failure mechanism analysis of slow-moving reservoir landslide", *J. Hydrol.*, **628**, 130588. https://doi.org/10.1016/j.jhydrol.2023.130588

Song, X. and C. Liu (2021), "Experimental study of the floor-attached vortices in pump sump using V3V", *Renew. Energy*, **164**, 752-766. https://doi.org/10.1016/j.renene.2020.09.088

Sun, T., F. Chen, L. Zhong, W. Liu and Y. Wang (2019), "GIS-based mineral prospectivity mapping using machine learning methods: A case study from Tongling ore district, eastern China", *Ore Geol. Rev.*, **109**, 26-49. https://doi.org/10.1016/j.oregeorev.2019.04.003

Tamayo-Trujillo, R., V.A. Ruiz-Pozo, S. Cadena-Ullauri, P. Guevara-Ramírez, E. Paz-Cruz, R. Zambrano-Villacres, D. Simancas-Racines and A.K. Zambrano (2024), "Molecular mechanisms of semaglutide and liraglutide as a therapeutic option for obesity", *Front. Nutr.*, **11**, 1398059. https://doi.org/10.3389/fnut.2024.1398059

Toghroli, A., M. Mohammadhassani, M. Suhatril, M. Shariati and Z. Ibrahim (2014), "Prediction of shear capacity of channel shear connectors using the ANFIS model", *Steel Compos Struct.*, **17**(5), 623-639. http://doi.org/10.12989/scs.2014.17.5.000

Tran, L.M., A. Santoro, L. Liu, S.A. Josselyn, B.A. Richards and P.W. Frankland (2022), "Adult neurogenesis acts as a neural regularizer", *Proceedings of the National Academy of Sciences*, **119**(45), e2206704119. https://doi.org/10.1073/pnas.2206704119

Trimm, E. and K. Red-Horse (2023), "Vascular endothelial cell development and diversity", *Nature Rev. Cardiol.*, **20**(3), 197-210. https://doi.org/10.1038/s41569-022-00770-1

Trung, N.T., N. Alemi, J.H. Haido, M. Shariati, S. Baradaran and S.T. Yousif (2019), "Reduction of cement consumption by producing smart green concretes with natural zeolites", *Smart Struct. Syst.*, **24**(3), 415-425. https://doi.org/10.12989/sss.2019.24.3.415

Verde, L., E. Frias-Toral, S. Cacciapuoti, D. Simancas-Racines, M. Megna, G. Caiazzo, L. Potestio, M. Maisto, G. C. Tenore and A. Colao (2024), "Very low-calorie ketogenic diet (VLCKD), a therapeutic nutritional tool for acne?", *J. Transl. Med.*, **22**(1), 322. https://doi.org/10.1186/s12967-024-05119-5

Verdugo-Paiva, F., P. Zambrano-Achig, D. Simancas-Racines and A. Viteri-García (2020), "Selective removal compared to complete removal for deep carious lesions", *Medwave*, **20**(1). https://doi.org/10.5867/medwave.2020.01.7758

Viteri-García, A., N. Montero, I. Arévalo-Rodríguez, A. Armas-Vega, C. Félix and D. Simancas-Racines (2018), "Odontología basada en evidencia: conceptos generales y su relevancia", Revista Kiru **15**(1). https://doi.org/10.24265/kiru.2018.v15n1.06

Waclawovsky, G., M. L. Pedralli, B. Eibel, M. I. Schaun and A. M. Lehnen (2021), "Effects of different types of exercise training on endothelial function in prehypertensive and hypertensive individuals: A systematic review", *Arquivos Brasileiros de Cardiologia*, **116**, 938-947. https://doi.org/10.36660/abc.20190807

Wong, P.F., J. Jamal, K.L. Tong, E.S. Khor, C.E. Yeap, H.L. Jong, S.T. Lee, M.R. Mustafa and S. Abubakar (2017), "Deregulation of hsa-miR-20b expression in TNF-α-induced premature senescence of human pulmonary microvascular endothelial cells", *Microvasc. Res.*, **114**, 26-33. https://doi.org/10.1016/j.mvr.2017.06.002

Yang, H., C. He, Y. Bi, X. Zhu, D. Deng, T. Ran and X. Ji (2022), "Synergistic effect of VEGF and SDF-1α in endothelial progenitor cells and vascular smooth muscle cells", *Front. Pharmacol.*, **13**, 914347. https://doi.org/10.3389/fphar.2022.914347

Yang, H., G. Huang, C. Chen, Y. Yang, Q. Wang and X. Dai (2024), "Method for evaluation of geological strength index of carbonate cliff rocks: Coupled hyperspectral–digital borehole image technique", *J. Rock Mech. Geotech. Eng.*, **16**(10), 4204-4215. https://doi.org/10.1016/j.jrmge.2024.08.013

Yang, H.Q., C.W. Chen and J.H. Ni (2023), "A hyperspectral evaluation approach for quantifying salt-induced weathering of sandstone", *Sci. Total Environ.*, **885**, 163886. https://doi.org/10.1016/j.scitotenv.2023.163886

Yang, M., H. Liang, X. Wu and Z. Zhang (2025), "A flexible and efficient algorithm for high dimensional support vector regression", *Neurocomputing*, **611**, 128671. https://doi.org/10.1016/j.neucom.2024.128671

Yazdani, M., K. Kabirifar, B.E. Frimpong, M. Shariati, M. Mirmozaffari and A. Boskabadi (2021), "Improving construction and demolition waste collection service in an urban area using a simheuristic approach: A case study in Sydney, Australia", *J. Clean. Prod.*, **280**. https://doi.org/10.1016/j.jclepro.2020.124138

Yazdani, M., K. Kabirifar, B. E. Frimpong, M. Shariati, M. Mirmozaffari and A. Boskabadi (2021), "Improving construction and demolition waste collection service in an urban area using a simheuristic approach: A case study in Sydney, Australia", Journal of Cleaner Production **280**: 124138.doi:

Zambrano-Sánchez, R., P. Alvarez-Mena, D. Hidalgo, C.E. Liquitay, J.V. Franco, R.W. Vernooij, D. Simancas-Racines, A. Viteri-García and C. Montesinos-Guevara (2022), "Quality assessment of clinical practice guidelines (CPG) for the diagnosis and treatment of inflammatory bowel disease using the AGREE II instrument: A systematic review", BMC gastroenterology **22**(1), 447. https://doi.org/10.1186/s12876-022-02539-9

Zambrano, A.K., S. Cadena-Ullauri, P. Guevara-Ramírez, E. Frias-Toral, V.A. Ruiz-Pozo, E. Paz-Cruz, R. Tamayo-Trujillo, S. Chapela, M. Montalván and G. Sarno (2023), "The impact of a very-low-calorie ketogenic diet in the gut microbiota composition in obesity", *Nutrients*, **15**(12), 2728. https://doi.org/10.3390/nu15122728

Zambrano, A. K., S. Cadena-Ullauri, V. A. Ruiz-Pozo, R. Tamayo-Trujillo, E. Paz-Cruz, P. Guevara-Ramírez, E. Frias-Toral and D. Simancas-Racines (2024), "Impact of fundamental components of the Mediterranean diet on the microbiota composition in blood pressure regulation", *J. Transl. Med.*, **22**(1), 417. https://doi.org/10.1186/s12967-024-05175-x

Zambrano, A.K., E. Paz-Cruz, V.A. Ruiz-Pozo, S. Cadena-Ullauri, R. Tamayo-Trujillo, P. Guevara-Ramírez, R. Zambrano-Villacres and D. Simancas-Racines (2024), "Microbiota dynamics preceding bariatric surgery as obesity treatment: A comprehensive review", *Front. Nutr.*, **11**, 1393182. https://doi.org/10.3389/fnut.2024.1393182

Zandi, Y., M. Shariati, A. Marto, X. Wei, Z. Karaca, D. K. Dao, A. Toghroli, M. H. Hashemi, Y. Sedghi and K. Wakil (2018), "Computational investigation of the comparative analysis of cylindrical barns subjected to earthquake", *Steel Compos. Struct.*, **28**(4), 439-447. http://doi.org/10.12989/scs.2018.28.4.439

Zhang, B., M. Yuan, L. Cheng, W. Jiao and C. Luo (2024), "Exploring the vortex structure and its dynamic behavior induced by a closed pump sump", *Ocean Eng.*, **307**, 118103. https://doi.org/10.1016/j.oceaneng.2024.118103

Zhao, X., X. Zhang, S. Zheng and Q. Yang (2025), "Residual seismic performance and fragility assessment of corroded steel beam-column welded joints", *Structures*, **73**, 108455. https://doi.org/10.1016/j.istruc.2025.108455

Zhou, Y., Y. Li and C. Chen (2024), "The key role of digital governance, natural resource depletion, and industrialization in social well-being: A case study of China", *Resour. Policy*, **93**, 104969. https://doi.org/10.1016/j.resourpol.2024.104969

CC

Nomenclature

Cyclic Shear Viscosity	CSV
Support Vector Machines	SVMs
Nitric Oxide	NO

Flow-Mediated Dilation	FMD
High-Intensity Interval Training	HIIT
Resistance Exercise	RE
Low-Density Lipoprotein	LDL
Endothelial Cells	ECs
Endothelial Nitric Oxide Synthase	ENOS
Vascular Smooth Muscle Cell	VSMC
Machine Learning	ML
Convolutional Neural Networks	CNNs
Water Contact Angle	WCA
Electrochemical Corrosion Potential	Ecorr
Rectified Linear Unit	ReLU
Potentio Dynamic Polarization	PDP
Adaptive Gradient Algorithm	AdaGrad
Radial Basis Function	RBF
Mean Absolute Error	MAE
Mean Squared Error	MSE
Shapley Additive exPlanations	SHAP
Local Interpretable Model-agnostic Explanations	LIME
True Positives	TP
True Negatives	TN
False Positives	FP
False Negatives	FN
Corrosion Current Density	icorr
Random Forests	RF
Standard Deviation	SD