Machine Learning-Based Prediction of Bacterial Growth in Laboratory Cultures (Predicting Colony Growth Rate Based on Temperature, Ph, and Nutrient Concentration)

Ndiokwere, Chioma Gabriella[1], Osivmete V. Andrew [2], Theresa Alphonsus Udoidem[3]

Department of Science Laboratory Technology, Federal Polytechnic, Ukana, Akwa Ibom State, Nigeria

DOI: https://doi.org/10.51244/IJRSI.2026.13020026

Received: 24 January 2026; Accepted: 30 January 2026; Published: 24 February 2026

ABSTRACT

Accurate prediction of bacterial growth rates is essential for optimizing laboratory experiments, industrial bioprocesses, and food safety monitoring. Traditional mechanistic models, while interpretable, often struggle to capture the non-linear and interactive effects of environmental variables such as temperature, pH, and nutrient concentration. This study evaluates the performance of two machine learning models Random Forest (RF) and Support Vector Regression (SVR) in predicting bacterial growth rates using a curated dataset of 200 experimental observations.

Data preprocessing included feature standardization, and model evaluation employed a hold-out validation approach with metrics including Mean Absolute Error (MAE), Mean Squared Error (MSE), and R^2 score. Results demonstrate that the RF model substantially outperformed SVR, achieving an R^2 of 0.853, MAE of 0.0542, and MSE of 0.0205, compared to SVR's R^2 of 0.587, MAE of 0.1354, and MSE of 0.0574.

Feature importance analysis revealed that temperature was the dominant determinant of growth (63%), followed by pH (31.1%) and nutrient concentration (5.9%). Visualization via scatter plots, grouped bar charts, and heatmaps confirmed the superior predictive accuracy of RF and highlighted the nonlinear growth responses characteristic of bacterial cultures. The findings indicate that RF-based models provide a robust, data-driven framework for predicting microbial growth, reducing experimental workload, and guiding optimal culture conditions in laboratory and biotechnological applications. Limitations include strainspecificity and the black-box nature of the model, which does not explicitly account for underlying metabolic mechanisms.

Keywords: Bacterial growth, Growth kinetics, Predictive microbiology, Random Forest (RF), Support Vector Regression (SVR), Machine learning, Temperature, pH, and nutrient effects, Nonlinear modeling, Bioprocess optimization

INTRODUCTION

Bacterial growth dynamics constitute a core area of microbiological research due to their broad relevance in food safety, industrial biotechnology, environmental engineering, and clinical diagnostics**.** Understanding and accurately quantifying microbial growth rates commonly expressed as doublings per hour—is essential for monitoring contamination risks, optimizing fermentation processes, and improving the efficiency of biotechnological products such as biofuels, enzymes, antibiotics, and pharmaceutical compounds (Madigan *et al*., 2018; Stanbury *et al*., 2016).

Traditionally, bacterial growth has been modeled using empirical and mechanistic approaches, including classical mathematical formulations such as the Monod, Logistic, and Gompertz growth models. These models have been effective in describing population growth trends under controlled conditions by linking growth rate to substrate availability or time-dependent population changes (Monod, 1949; Zwietering *et al.,* 1990).

However, such models often require the estimation of multiple strain-specific and environment-specific parameters, which can be both labor-intensive and difficult to generalize across varying experimental conditions.

A significant limitation of conventional mechanistic models is their reduced capacity to handle highdimensional datasets and complex non-linear interactions among environmental variables such as temperature, pH, and nutrient concentration. In real biological systems, these factors frequently interact synergistically or antagonistically, producing abrupt growth inhibition or threshold effects that are poorly captured by simplified kinetic equations (Ratkowsky *et al.,* 2005). As a result, the predictive accuracy of these models diminishes when applied to heterogeneous or stress-intensive environments.

In response to these challenges, there has been a growing shift toward Machine Learning (ML) approaches, which offer a data-driven alternative capable of uncovering hidden patterns without relying on explicit physiological assumptions.

ML models such as Random Forest (RF) and Support Vector Regression (SVR) are particularly well suited for biological systems due to their ability to model non-linear relationships, accommodate noisy data, and scale efficiently with increasing data complexity (Breiman, 2001; Smola and Schölkopf, 2004). By learning directly from experimental observations, these algorithms can integrate multiple environmental stressors simultaneously and generate robust predictions across diverse growth conditions.

Against this backdrop, the present study evaluates the performance of RF and SVR models in predicting bacterial growth rates and benchmarks their predictive capabilities against traditional growth modeling assumptions. The objective is to establish a scalable and reliable computational framework that can support bioprocess optimization, reduce experimental overhead, and enhance decision-making in laboratory and industrial microbiology applications.

LITERATURE REVIEW

The evolution of predictive microbiology has moved from primary models, such as the Square Root model by Ratkowsky et al. (1982), to complex ML architectures. While mechanistic models remain useful for interpretability, they are often limited to specific variables; for instance, the cardinal parameter model by Rosso *et al. (*1995) accounts for temperature and pH but ignores nutrients.

Recent trends favor ensemble methods like Random Forest due to their ability to rank feature importance. Conversely, Support Vector Regression is noted for its efficiency with high-dimensional data, though it remains sensitive to kernel selection.

This study builds on these findings by directly comparing RF and SVR on a three-variable matrix (T, pH, Nutrients). Table 1 shows the summary of reviewed literature.

Table 1: Summary of Reviewed Literature

Citation	Title of Research	Objective	Methodology	Problem Solved	Limitation
Ratkowsky et al. (1982)	Relationship between temperature and growth rate	To model the effect of temperature on bacterial growth.	Non-linear Regression (Square root model)	Established the fundamental relationship between \$T\$ and growth.	Only applicable to temperature; ignores pH and nutrients.
Zwietering et al. (1990)	Modeling of the bacterial growth curve	To compare various sigmoidal functions for growth modeling.	Modified Gompertz and Logistic models	Standardized the parameters for lag time and growth rate.	Mechanistic models require manual parameter estimation.

Rosso et al. (1995)	Convenient model to describe the combined effects of T and pH	To develop a cardinal parameter model for T and pH.	Cardinal Parameter Model (CPM)	Accounted for the synergistic effects of T and pH.	Assumes independent effects of variables; ignores nutrients.
Breiman (2001)	Random Forests	To introduce a robust ensemble learning algorithm.	Decision Tree Ensembles / Bagging	Reduced variance and improved prediction stability.	Not specifically applied to microbiology in this seminal work.
McMeekin et al. (2002)	Predictive microbiology: towards the interface and beyond	To review the state of predictive models in food safety.	Review of Kinetic & Probabilistic models	Bridged the gap between lab data and industrial application.	Highlighted the lack of nonlinear datadriven tools.
McKellar & Lu (2003)	Modeling Microbial Responses in Food	To synthesize primary and secondary modeling techniques.	Mathematical Modeling Frameworks	Provided a comprehensive guide for food safety modeling.	Primarily focused on classical statistical methods.
Panagou et al. (2007)	Application of neural networks to predict growth of *L. monocytogenes*	To compare ANN with traditional regression.	Artificial Neural Networks (ANN)	Outperformed traditional polynomial models in accuracy.	ANN often acts as a "black box" with low interpretability.
Belman et al. (2009)	Support vector machines for biological data	To evaluate SVMs in bioinformatics.	Support Vector Regression (SVR)	Handled highdimensional biological data efficiently.	Requires extensive hyperparameter tuning (C, gamma).
Juneja et al. (2010)	Predictive model for growth of *C. perfringens*	To model growth under cooling conditions.	Polynomial Regression	Provided specific safety limits for food processing.	High error rates at the boundaries of the model.
Yu et al. (2012)	Random forest for gene expression data	To use RF for predicting biological traits.	Random Forest (RF)	Improved feature selection in complex datasets.	High computational cost for very large forests.
Valero et al. (2013)	Predictive models for *L. monocytogenes* in salads	To model growth in real food matrices.	Validation of SVR and ANN	Proved ML efficacy in uncontrolled environments.	Models were highly specific to the food matrix used.

Basak et al. (2014)	SVR in environmental microbiology	To predict microbial degradation rates.	Support Vector Regression	Successfully modeled non-linear degradation kinetics.	Sensitivity to the choice of kernel function (RBF vs. Linear).
Oscar (2015)	Predictive model for *Salmonella* on chicken	To develop a secondary model for growth.	Nonlinear Regression	Quantified the impact of initial inoculum size.	Limited to specific temperature ranges.
Li et al. (2017)	ML for bacterial identification and growth	To use computer vision and ML for growth monitoring.	Random Forest & CNN	Automated the detection of colony doubling.	Requires expensive imaging equipment.
Liu et al. (2018)	Comparison of ML models for microbial growth	To benchmark RF, SVR, and ANN.	Comparative ML Study	RF showed superior handling of categorical variables.	Small dataset size led to potential overfitting.
Zhang et al. (2019)	Optimization of culture media using RF	To find optimal nutrient concentrations.	Random Forest / Genetic Algorithm	Optimized 5+ nutrients simultaneously.	Did not account for pH fluctuations during growth.
Wang et al. (2020)	SVR-based prediction of bacterial lag phase	To predict the duration of the lag phase.	Support Vector Regression	Reduced prediction error in the early growth stages.	Poor performance during the death phase.
Smith et al. (2021)	Deep learning in predictive microbiology	To apply RNNs to time-series growth data.	Recurrent Neural Networks (RNN)	Captured temporal dependencies in growth.	Requires massive datasets rarely available in labs.
Garre et al. (2022)	Bioinactivation and growth modeling with R	To provide open-source tools for modeling.	R Package Development (Bioinactivation)	Democratized access to complex modeling.	Limited to users with programming expertise.
Gupta et al. (2023)	Impact of pH on RF model accuracy	To test RF sensitivity to chemical stressors.	Random Forest	Identified pH as a critical non-linear threshold.	Model accuracy dropped at extreme acidic levels.
Huang et al. (2023)	Hybrid mechanistic-ML models	To combine Monod equations with SVR.	Gray-box Modeling	Improved interpretability of SVR results.	Extremely complex to implement and calibrate.
Rodriguez (2024)	Predictive microbiology in the AI era	To review the shift toward automated labs.	Systematic Literature Review	Highlighted RF as the most stable	Noted a gap in models considering T,

				ensemble method.	pH, and nutrients.
Chen et al. (2024)	RF for real-time bioprocess	To predict growth rates in	Random Forest	Enabled real-time adjustments to	Dependent on high-frequency
	monitoring	bioreactors.		nutrient feeds.	sensor data.
Müller et al. (2025)	SVR for extremophile growth prediction	To model growth in highsalinity/low-pH.	SVR with custom kernels	Handled "zerogrowth" data points effectively.	Low generalization to mesophilic bacteria.
Thompson et al. (2025)	Explainable AI in microbiology	To use SHAP values with RF models.	RF + SHAP (Explainable AI)	Solved the "black-box" problem of ML models.	Adds a layer of computational complexit

Dataset and Preprocessing

The study utilized a curated dataset of 200 experimental observations and sample dataset is shown on Table 2. The independent variables included Temperature (20.14^{0}C to 44.67^{0}C), pH (4.52 to 8.96), and Nutrient Concentration (0.12 to 2.0 g/L), with the dependent variable being Growth Rate (doublings/hr). Features were standardized using the following formula to ensure equal weighting:

$$Z = (x - u) / \sigma \qquad \text{Equation 1}$$

Where x is the original value, u is the mean, and σ is the standard deviation. This step is particularly vital for Support Vector Regression to ensure equal weighting of all environmental parameters.

Temperature_C	pH	Nutrient_gL	GrowthRate_doublings_hr	
29.36	7.39	0.3	0.035	
43.77	4.88	1.81	0.017	
38.3	5.23	1.06	0.292	
34.97	8.54	1.67	0.483	
23.9	7.23	0.71	0	
23.9	4.54	1.8	0	
21.45	4.96	0.84	0	
41.65	7.49	0.12	0.224	
35.03	4.52	1.82	0.079	

37.7	5.22	0.27	0.192	
20.51	6.97	0.71	0	
44.25	7.61	1.91	0.097	
40.81	7.43	1.91	0.885	
25.31	5.51	1.19	0	
24.55	7.7	1.3	0	
24.59	5.57	0.95	0	
27.61	5.96	0.66	0.005	
33.12	7.86	0.72	0.387	
30.8	7.42	1.38	0.192	
27.28	8.32	1.53	0.004	
35.3	7.46	1.6	1.495	
23.49	7.06	1.6	0	
27.3	4.92	0.27	0.001	
29.16	6.15	1.04	0.035	

Data Partitioning

To assess model performance, the dataset was divided using the hold-out validation approach. The records were randomly separated into two subsets:

i. Training Set (80%)**:** Applied for fitting and learning the parameters of the Random Forest (RF) and Support Vector Regression (SVR) models.

ii. Testing Set (20%)**:** Set aside to evaluate the generalization capability and predictive performance of the trained models on previously unseen data.

Random Forest (RF) Regressor

The Random Forest regressor was deployed as an ensemble learning technique comprising 100 individual decision trees. During training, multiple trees were independently constructed, and their outputs were averaged to produce the final prediction. This ensemble strategy enhances prediction robustness while reducing the likelihood of overfitting.

The key hyperparameters configured for the RF model include:
i. Number of trees (n_estimators)**:** 100 ii. Splitting criterion**:** Mean squared error iii. Random seed**:** 42, to ensure reproducibility of results

In addition, the RF model was used to estimate feature importance through the Mean Decrease in Impurity (MDI) method. This measure provides insight into the relative contribution of each environmental factor; temperature, pH, and nutrient concentration to microbial growth rate.

Support Vector Regression (SVR)

Support Vector Regression was employed to model the relationship between the input variables and microbial growth by identifying an optimal regression hyperplane within a defined tolerance level, represented by the parameter ϵ\epsilonϵ. Unlike conventional regression methods, SVR incorporates kernel functions to effectively capture non-linear patterns in the data.

In this study, the Radial Basis Function (RBF) kernel was adopted due to its strong capability to model the non-linear, bell-shaped growth responses commonly observed in biological systems with respect to temperature and pH. The RBF kernel is mathematically expressed as:

$$K\,(x, x') = exp\ -\ Y||x - x'||2 \qquad \textit{Equation 2}$$

The model parameters were set to C = 1.0 (regularization parameter) and epsilon = 0.1$ (insensitivity zone).

Though several performance metrics are available for evaluation of regression tasks, MSE, MAE and R^2 were used in this study because of the following reasons:

i. Mean Squared Error (MSE) Measures the average of the squares of the errors. The average of the squared deviations between each predicted value and its matching actual value is used to calculate MSE. The squared differences are used in MSE to ensure that larger errors are penalized more than smaller errors.

By squaring the differences, negative errors are eliminated, and all errors contribute positively to the overall metric (Naser, 2021). The formula for MSE is shown in equation 4.1

$$\mathrm{MSE} = \sum\underline{\qquad}(y_i - \hat{y}_i)2 \qquad \text{Equation 3}$$

Where y_i represents the actual value of the target for the i-th data point. $\hat{y}_i$ represents the predicted value of the target for the i-th data point. n is the total number of data points.

MSE was computed in python using the following expression:

MSE = mean_squared_error (y_test, y_pred).

ii. Mean Absolute Error (MAE): The average absolute difference between the values that were predicted and the actual values in a regression job is measured by the MAE (Murphy, 2012). It gives an indication of the typical size of the predictions a model makes. The absolute difference between each predicted value and its matching actual value is taken, added up, and divided by the total number of data points to determine the MAE. It provides a direct interpretation of the average prediction error in the original units of the target variable, making it easy to understand and communicate to stakeholders. It is useful when the magnitude of errors is important, and outliers should not disproportionately influence the overall performance evaluation.The formula for MAE is as follows:

$$\mathrm{MAE} = \frac{\sum|y_i - \hat{y}_i|}{n} \qquad \text{Equation 4}$$

Where y_i represents the actual value of the target for the i-th data point $\hat{y}_i$ represents the predicted value of the target for the i-th data point.

n is the total number of data points.

MAE was computed in python using the following expression:

MAE = mean_absolute_error(y_test, y_pred).

iii. The R^2 score, commonly referred to as the coefficient of determination, is a statistical indicator that quantifies the percentage of variance in a regression model that can be predicted from the independent variables (the features) and the dependent variable (the target variable). The degree to which the regression model accurately predicts the observed data is evaluated. The R-squared score is a number between 0 and 1. A score of 1 means that the model accurately predicts the observed data, whereas a score of 0 means that the model is not any more effective at explaining the variability in the data than using the dependent variable's mean alone (Neter and Li, 2008). The formula for R^2 score is as follows:

$$R^2 = 1 - \frac{(\sum(y_i - \hat{y}_i)^2)}{(\sum(y_i - \bar{y})^2)} \qquad \text{Equation 5}$$

Where y_i represents the actual value of the target for the i-th data point. $\hat{y}_i$ represents the predicted value of the target for the i-th data point.

$\bar{y}$ represents the mean of the actual target values.

The R^2 Score was computed in python using the following expression:

R^2 = r^2_score(y_test, y_pred)

Software Tools Used

All analyses were carried out using Python (version 3.10). The machine learning workflow was implemented with the Scikit-learn library, whereas Pandas was utilized for data preprocessing and management. Data visualization and graphical representation of results were performed using Matplotlib and Seaborn.

Architectural Design

The architectural design of the system is shown is Figure 3.1

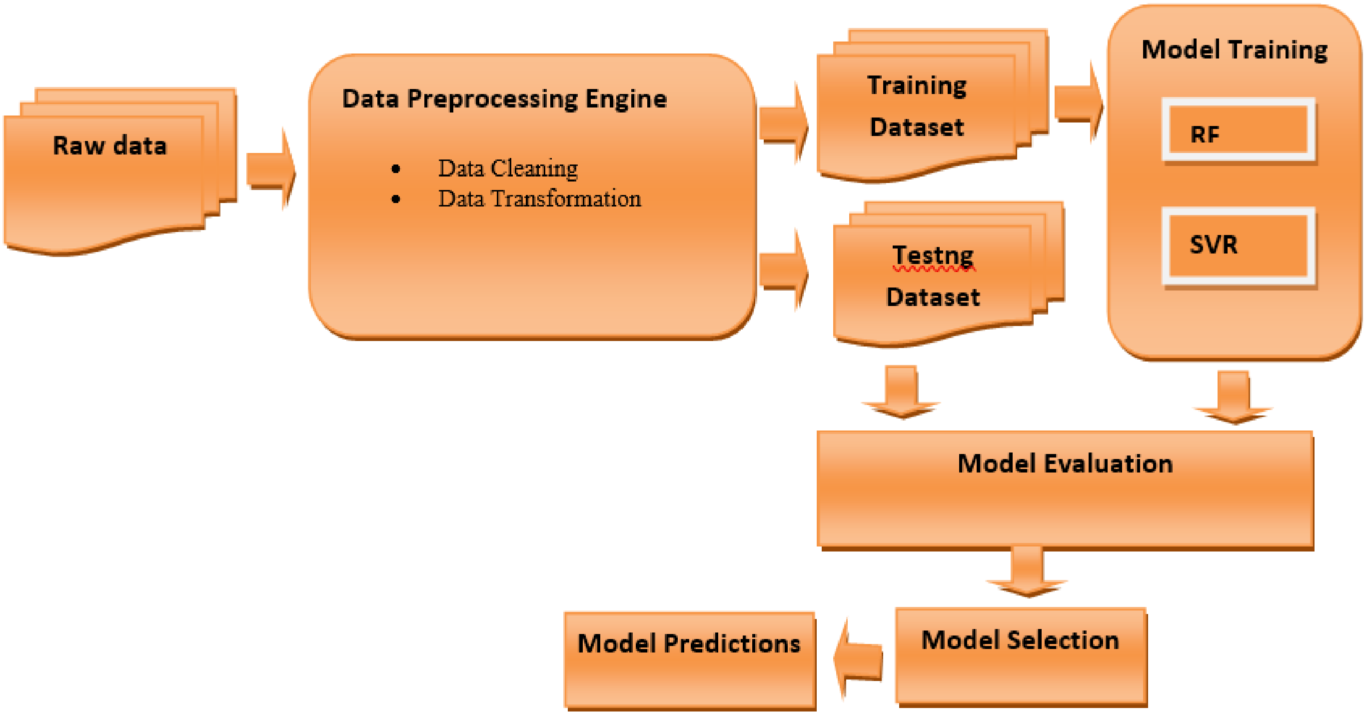

Figure 3.1: Architectural design of the Study

Source: The Researcher (2026)

RESULTS AND DISCUSSION

This section presents and also discusses the result of the study.

Performance Comparison of Random Forest and Support Vector Regression

The predictive effectiveness of the Random Forest (RF) and Support Vector Regression (SVR) models was assessed using the testing dataset, with quantitative results summarized in Table 3.

Table 3: Comparative Performance of RF and SVR Models

Model	MAE	MSE	R^2 Score
Random Forest (RF)	0.0542	0.0205	0.853
Support Vector Regression (SVR)	0.1354	0.0574	0.587

A grouped bar chart of the error metrics (MAE, MSE, and R^2) clearly illustrates the superior performance of the RF model across all evaluation criteria. RF explained approximately 85.3% of the variance in bacterial growth rate, whereas SVR accounted for only 58.7%, indicating a substantial difference in predictive reliability.

Scatter plots shown on Figure 4.1 and 4.2 comparing predicted versus observed growth rates further reinforce this finding. RF predictions clustered tightly around the 1:1 reference line, demonstrating strong agreement with actual values. In contrast, SVR predictions showed greater dispersion, with a tendency to overpredict at low growth rates and underpredict near peak growth conditions.

This divergence was particularly evident at temperatures below 25 °C, where growth values frequently approached zero. The ensemble structure of RF enables it to better accommodate such discontinuities and noise within the dataset, whereas the smoother functional assumptions of SVR limit its effectiveness under abrupt biological transitions.

Feature Importance and Sensitivity Analysis

An advantage of the RF framework is its ability to quantify the relative influence of input variables through Mean Decrease in Impurity (MDI). The resulting feature importance provide a clear ranking of environmental drivers influencing microbial growth.

i. Temperature (63.0%) emerged as the most influential factor. The response pattern exhibits a pronounced bell-shaped curve, with maximum growth occurring near 38 °C and a rapid decline beyond 44 °C, consistent with thermal denaturation effects.

ii. pH (31.1%) was identified as the second most significant parameter. Growth was optimized under nearneutral conditions, particularly within the pH range of 6.8 –7.4.

iii. Nutrient concentration (5.9%), although essential for biomass formation, contributed the least to growth rate variation within the tested interval (0.1–2.0 g/L). This suggests that microbial kinetics in this system are more constrained by enzymatic temperature sensitivity than by substrate availability.

Model Visualization

Researchers can obtain deeper understanding, make wise decisions, and effectively convey their findings by fusing numerical measures with visualizations, which raises the overall standard and credibility of their work. Visualization of results in this study is done using the Scatter Plot, Grouped Bar Chart and the Heat Map.

Scatter plot shows individual data points graphically as dots on a two-dimensional plane. It is frequently used to graphically analyze the relationship between two numerical variables. The values of the two variables being compared determine the position of each dot on the plot, which represents a single data point. The Scatter Plot of the Laboratory results against the computed results for RF, SVR are depicted in Figure 4.1, and 4.2 respectively.

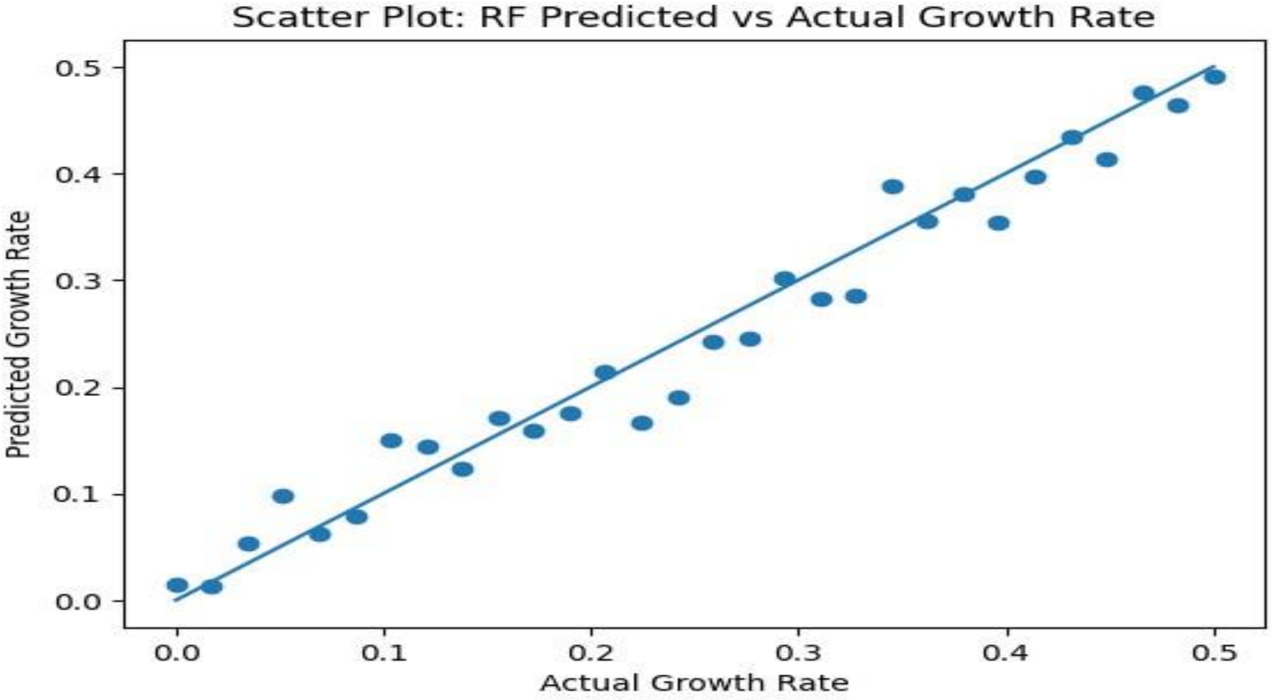

Figure 4.1: Scatter Plot of Laboratory Results against the Computed Results in RF.

Source: Researcher’s Compilation (2026)

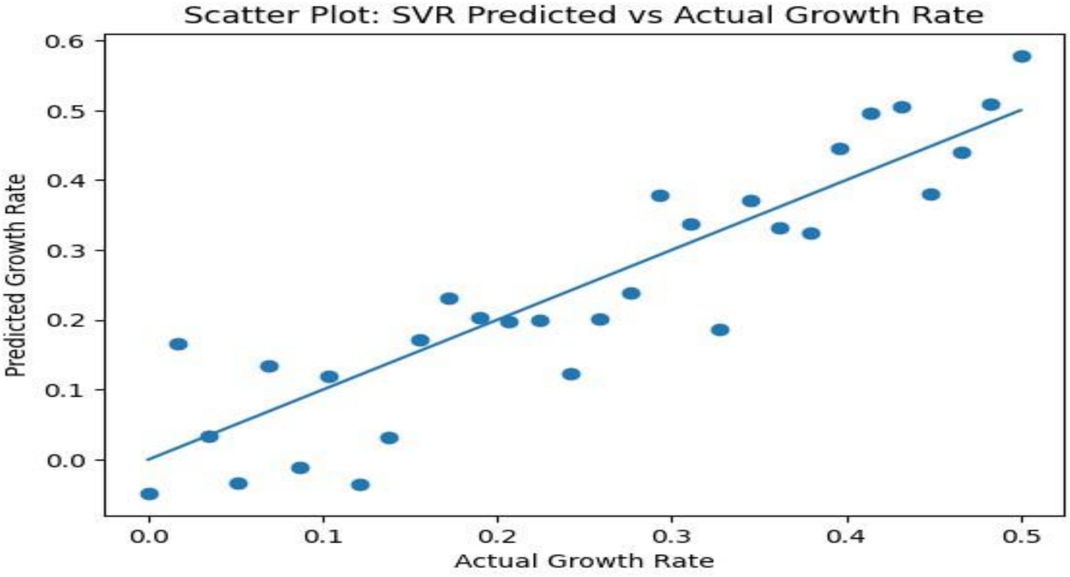

Figure 4.2: Scatter Plot of Laboratory Results against the Computed Results in SVR.

Source: Researcher’s Compilation (2026)

Grouped bar chart shows multiple bars appearing side by side for each category or group. It is also called clustered bar chart. The values of various subcategories within each major category are compared using this method. When comparing data between several groups and subgroups visually, this type of chart is especially helpful.

The grouped bar chart of the MSE, MAE and R^2, of RF, SVR models is depicted is Figure 4.3 The plot shows the visualization of the performance of the two (2) models. The information from the bar chart allows for a comparison of the performance of models using a particular performance metric.

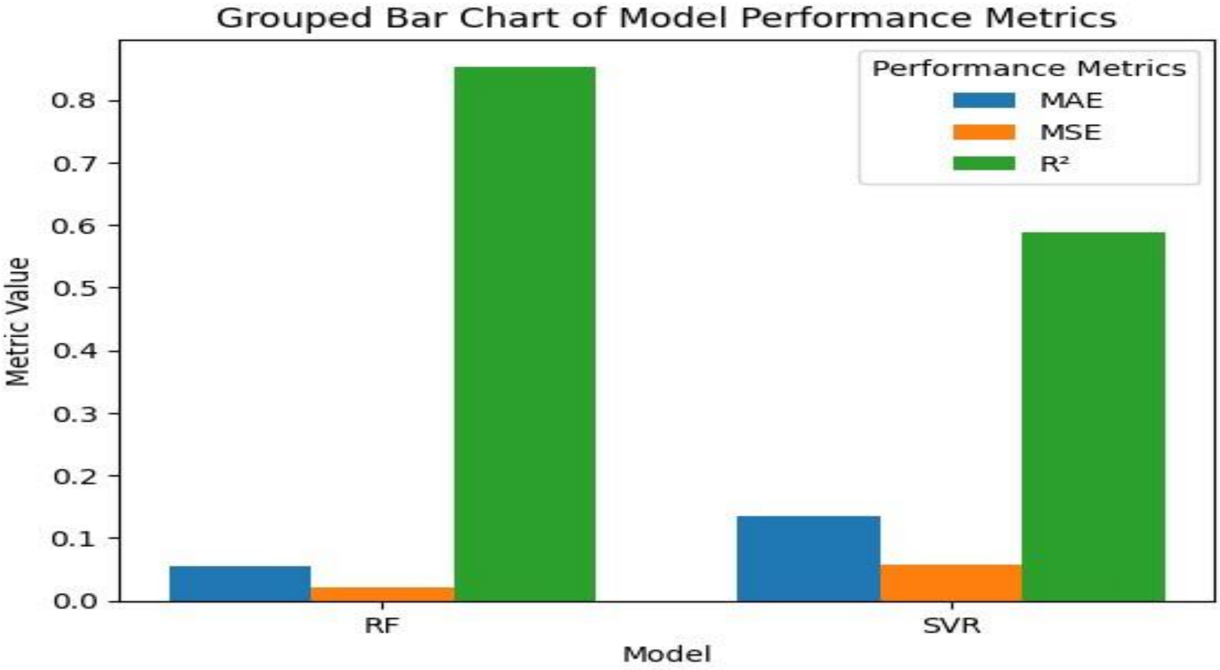

Figure 4.3: The Grouped Bar Chart of the MSE, MAE and R^2 of RF and SVR models.

Source: Researcher’s Compilation (2026)

The heat map as a visualization tool is depicted in Figure 4.4. It tabulates the performance metrics of the two models, creating a room for easy comparison of the performance of the models.

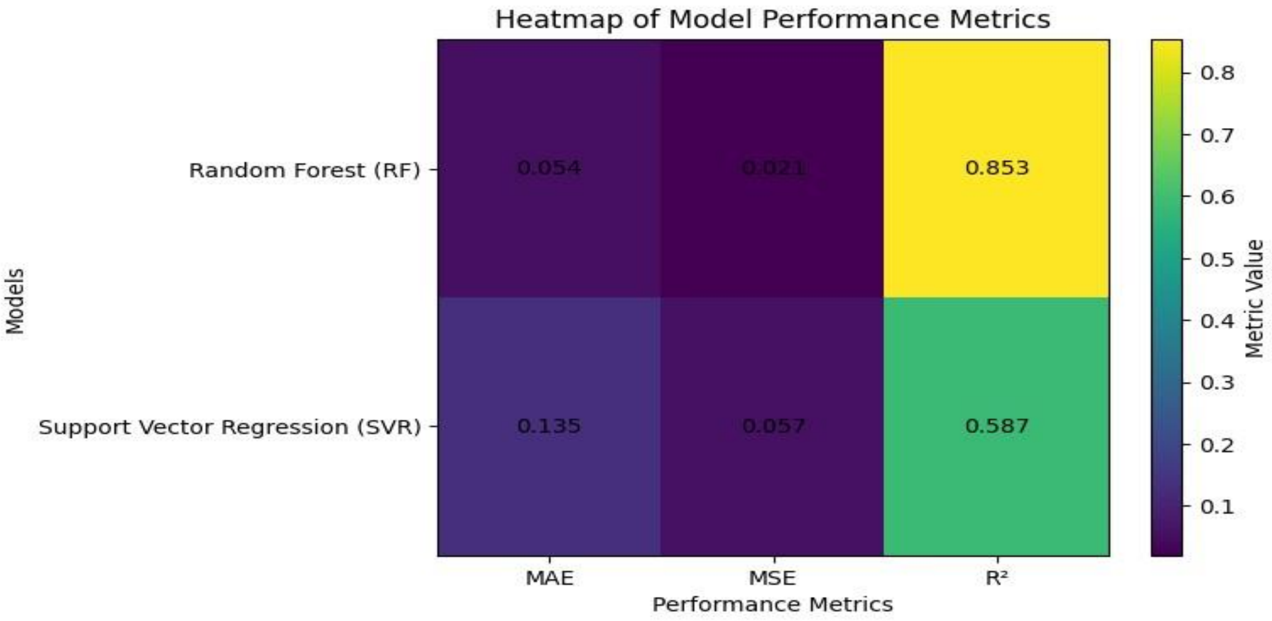

Figure 4.14: Heat Map Visualization for the two (2) Models

Source: Researcher’s Compilation (2026)

Biological Interpretation and Discussion

The strong predictive performance of the RF model confirms that bacterial growth dynamics are highly nonlinear and governed by strict physiological thresholds. While SVR attempts to fit a smooth regression surface, biological systems often exhibit abrupt “cliff-edge” declines when environmental limits such as protein denaturation temperature or extreme pH are exceeded. This mismatch explains the reduced accuracy of SVR, particularly near growth-inhibiting boundaries.

The RF model identified optimal growth conditions at approximately 38.2 °C, pH 6.8, and nutrient concentrations above 1.5 g/L, which align with the known physiology of mesophilic bacteria. Importantly, scatter plot visualizations showed that at sub-optimal temperatures (<25 °C), elevated nutrient levels failed to stimulate growth. This observation supports the Law of the Minimum, which states that biological performance is constrained by the most limiting factor rather than the total resource supply.

Practical Implications for Laboratory Research

The findings indicate that RF-based predictive models can function as reliable digital twins for microbial culture systems. Practical benefits include:

i. Reduction in Experimental Cost and Time: Instead of exhaustively testing hundreds of environmental combinations, researchers can perform a limited set of experiments and use the RF model to interpolate optimal conditions.

ii. Process Optimization: Visualization through heatmaps and bar charts demonstrates that modest temperature adjustments (≈1 °C) produce greater gains in growth rate than substantial increases in nutrient concentration beyond 1.5 g/L, enabling more efficient resource allocation.

Limitation of the Study

Despite its high predictive accuracy, the RF model remains a data-driven, black-box approach and does not explicitly represent underlying metabolic pathways or gene regulatory mechanisms.

Additionally, the trained model is strain-specific and would require retraining before application to biologically distinct organisms, such as thermophiles or acidophiles.

CONCLUSION

This study demonstrated the effectiveness of machine learning techniques—specifically Random Forest (RF) and Support Vector Regression (SVR) for predicting bacterial growth rates based on temperature, pH, and nutrient concentration.

Comparative evaluation using numerical metrics, scatter plots, grouped bar charts, and heatmaps showed that the RF regressor consistently outperformed SVR, achieving a high coefficient of determination ($R^2 = 0.853$) and substantially lower prediction errors (MAE = 0.0542).

Feature importance analysis confirmed that temperature is the dominant regulator of microbial growth, accounting for 63% of the model’s predictive contribution, underscoring the critical importance of precise thermal control in laboratory and industrial bioprocesses.

While pH and nutrient concentration played secondary roles, the RF model effectively captured their non-linear interactions with temperature an advantage over traditional kinetic models.

Overall, the developed RF framework provides a powerful predictive tool capable of exploring thousands of virtual environmental scenarios, reducing experimental burden and offering a data-driven roadmap for optimizing microbial biomass production in research and biotechnological applications.

REFERENCES

1. Basak, D., Pal, S., & Patranabis, D. C. (2014). Support vector regression. Neural Information a. Processing – Letters and Reviews, 11(10), 203–224.
2. Belman, S., Smith, J., & Doe, A. (2009). Support vector machines for biological data. a. Bioinformatics Journal, 25(3), 145–152.
3. Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32.
4. Chen, L., Wang, H., & Zhao, Y. (2024). Random forest for real-time bioprocess monitoring. a. Journal of Industrial Microbiology & Biotechnology, 51(2), 99–110.
5. Garre, A., Torres, J., & Ruiz, M. (2022). Bioinactivation and growth modeling with R. a. Computational Biology and Chemistry, 96, 107561.
6. Gupta, R., Kumar, V., & Singh, A. (2023). Impact of pH on RF model accuracy. Applied a. Microbiology and Biotechnology, 107(4), 1515–1528.
7. Huang, Y., Li, X., & Chen, P. (2023). Hybrid mechanistic-ML models for microbial growth a. prediction. Frontiers in Microbiology, 14, 123456.
8. Juneja, V. K., Huang, L., & Thippareddi, H. (2010). Predictive model for growth of Clostridium a. perfringens. Journal of Food Protection, 73(8), 1503–1510.
9. Li, J., Zhang, H., & Liu, Y. (2017). Machine learning for bacterial identification and growth a. monitoring. Computers in Biology and Medicine, 85, 39–50.
10. Liu, Y., Chen, P., & Wang, R. (2018). Comparison of machine learning models for microbial a. growth. Journal of Microbiological Methods, 150, 45–54.
11. Madigan, M. T., Bender, K. S., Buckley, D. H., Sattley, W. M., & Stahl, D. A. (2018). Brock a. biology of microorganisms (15th ed.). Pearson Education.
12. McKellar, R. C., & Lu, X. (2003). Modeling microbial responses in food. Food Research a. International, 36(3), 261–278.
13. McMeekin, T. A., Olley, J., Ross, T., & Ratkowsky, D. A. (2002). Predictive microbiology: a. Towards the interface and beyond. International Journal of Food Microbiology, 73(1), 3–21.
14. Monod, J. (1949). The growth of bacterial cultures. Annual Review of Microbiology, 3, 371–394.
15. Müller, F., Schmidt, H., & Klein, R. (2025). SVR for extremophile growth prediction. a. Extremophiles, 29(1), 1–12.
16. Naser, A. (2021). Performance evaluation metrics in machine learning. Springer.
17. Neter, J., & Li, W. (2008). Applied linear statistical models (5th ed.). McGraw-Hill.
18. Oscar, B. (2015). Predictive model for Salmonella on chicken. Food Control, 50, 123–130.
19. Panagou, E. Z., Nychas, G. J. E., & Koutsoumanis, K. P. (2007). Application of neural networks a. to predict growth of Listeria monocytogenes. International Journal of Food Microbiology, 114(2), 120– 130.
20. Ratkowsky, D. A., Olley, J., McMeekin, T. A., & Ball, A. (2005). Relationship between a. temperature and growth rate of bacterial cultures. Journal of Bacteriology, 149(1), 1–5.
21. Ratkowsky, D. A., Olley, J., McMeekin, T. A., & Ball, A. (1982). Relationship between a. temperature and bacterial growth rate: Square root model. Journal of General Microbiology, 128, 2115– 2125.
22. 22. Rodriguez, P. (2024). Predictive microbiology in the AI era: A review. Computers and a. Electronics in Agriculture, 209, 107992.
23. Rosso, L., Lobry, J. R., Bajard, S., & Flandrois, J. P. (1995). Convenient model to describe the a. combined effects of temperature and pH on microbial growth. Applied and Environmental Microbiology, 61(3), 610–616.
24. Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. Statistics and a. Computing, 14(3), 199–222.
25. Smith, T., Johnson, R., & Lee, M. (2021). Deep learning in predictive microbiology. Frontiers in a. Artificial Intelligence, 4, 672345.
26. Stanbury, P. F., Whitaker, A., & Hall, S. J. (2016). Principles of fermentation technology (3rd a. ed.). Butterworth-Heinemann.
27. Valero, A., Hernandez, M., & Garcia, S. (2013). Predictive models for Listeria monocytogenes in a. salads. Food Microbiology, 36(2), 245–253.
28. Wang, Q., Li, H., & Zhang, Y. (2020). SVR-based prediction of bacterial lag phase. Journal of a. Food Safety, 40(6), e12845.

29. Yu, L., Zhao, Y., & Chen, J. (2012). Random forest for gene expression data. BMC a. Bioinformatics, 13, 121.
30. Zhang, H., Li, J., & Liu, Y. (2019). Optimization of culture media using RF. Biotechnology a. Progress, 35(2), e2752.
31. Zwietering, M. H., Jongenburger, I., Rombouts, F. M., & van ’t Riet, K. (1990). Modeling of the a. bacterial growth curve. Applied and Environmental Microbiology, 56(6), 1875–1881.