RESEARCH PAPER

Machine-Learning-Driven Evaluation of Vetiveria zizanioides and Microbial Systems for Sustainable Dye Wastewater Treatment

Saju S[1], MohanaPriya J[2], Markco M[3]

[1] Associate Professor, Dept. of Biomedical Engineering, KIT-Kalaignarkarunanidhi Institute of Technology, Coimbatore, Tamilnadu, India. Email: Sajube82@gmail.com

[2] Associate Professor, Dept. of Biotechnology, KIT-Kalaignarkarunanidhi Institute of Technology, Coimbatore, Tamilnadu, India.

[3] Assistant Professor, Dept. of AI & DS, KIT-Kalaignarkarunanidhi Institute of Technology, Coimbatore, Tamilnadu, India.

How to cite this article: Saju S, MohanaPriya J, Markco M. Machine-Learning-Driven Evaluation of Vetiveria zizanioides and Microbial Systems for Sustainable Dye Wastewater Treatment. Int J Drug Deliv Technol. 2026;16(37s): 140-146. DOI: 10.25258/ijddt.16.37s.21

***Abstract*—The presence of synthetic dyes such as methyl orange in textile wastewater poses a serious environmental challenge due to their persistence and toxicity. Conventional treatment methods are energy-intensive and often lead to secondary pollution. In this study, Vetiveria zizanioides (Vetiver plant) and a microbial consortium were used as eco-friendly biological agents to remove methyl orange from contaminated water. To enhance predictive capability, Artificial Intelligence (AI) models were integrated with the experimental findings to estimate decolorization efficiency, aromatic amine removal, seed germination, and radicle growth. The Random Forest and XGBoost regression algorithms were trained using laboratory-reported datasets to predict treatment outcomes. The models achieved a high degree of correlation (R^2 = 0.98) between predicted and reported values, with a minimal mean absolute error of less than 2%. Results revealed that plant-based treatment achieved 93% dye decolorization and 91.2% aromatic amine removal, outperforming microbial and control samples. The AI predictions closely aligned with experimental values, demonstrating the feasibility of data-driven modeling in bioremediation. This integrated approach bridges environmental biotechnology and computational intelligence, enabling faster decision-making for large-scale wastewater treatment applications.**

Keywords—Vetiveria zizanioides, Bioremediation, Methyl Orange Dye, Predictive Modeling, Random Forest, XGBoost.

I. INTRODUCTION

Liver The textile and dyeing industries contribute significantly to global economic growth but also generate large volumes of colored effluents that pose a major threat to aquatic ecosystems. Among the various classes of synthetic dyes, azo dyes such as methyl orange are widely used because of their low cost, chemical stability, and bright coloration. However, these same properties make them highly resistant to degradation. It is estimated that nearly 10–15% of synthetic dyes used in dyeing processes are lost in wastewater, leading to serious water contamination problems. These dyes not only reduce sunlight penetration and oxygen transfer in aquatic systems but also release toxic aromatic amines during decomposition, which can be mutagenic, carcinogenic, and teratogenic to living organisms. The persistence of such compounds in industrial effluents calls for efficient, sustainable, and low-cost treatment technologies. Conventional physicochemical methods such as adsorption, coagulation–flocculation, electrochemical oxidation, ozonation, and advanced oxidation processes have been extensively applied for dye removal. Although these methods are effective, they are often limited by high operational costs, energy consumption, and generation of secondary sludge that requires additional disposal. In developing regions, particularly in rural and semi-urban textile clusters, these constraints make traditional treatment methods impractical. Hence, the focus has gradually shifted toward biological methods, which exploit the natural ability of plants and microorganisms to degrade or transform hazardous pollutants. Phytoremediation, the use of green plants to detoxify contaminated water and soil, has emerged as an eco-friendly and cost-effective approach. Among the available plant species, Vetiveria zizanioides (commonly known as vetiver grass) has gained attention for its exceptional adaptability, fast growth, and deep root system capable of absorbing and accumulating organic pollutants. Its rhizospheric region supports diverse microbial communities, and the oxygen release from roots enhances aerobic degradation of pollutants. Previous studies have shown that Vetiver can effectively remove heavy metals, hydrocarbons, and textile dyes from contaminated environments. Its use for methyl orange removal is particularly promising because the plant's fibrous root network can adsorb dye molecules while plant enzymes catalyze oxidative degradation. Parallel to plant-based systems, bioremediation using microbial consortia provides another natural strategy for dye degradation. Microorganisms such as bacteria, fungi, and actinomycetes produce enzymes like azoreductases, laccases, and peroxidases, which break the azo bonds (–N=N–) in dye molecules under anaerobic or facultative conditions. However, microbial systems can be sensitive to environmental parameters such as pH, temperature, oxygen concentration, and nutrient availability, which often limit their large-scale application. Despite these biological advances, a challenge remains: predicting and optimizing the performance of such systems under varying conditions. Biological treatment processes are inherently nonlinear and governed by complex interactions among physical, chemical, and biological parameters. Experimental studies, while informative, are time-consuming and resource-intensive, and they often provide limited insight into system dynamics. In this context, the integration of Artificial Intelligence (AI) and Machine Learning (ML) offers transformative potential for environmental monitoring and process optimization.AI-based predictive models can learn from experimental data to capture hidden patterns and relationships between input and output variables. Algorithms such as Random Forest and Extreme Gradient Boosting (XGBoost) are particularly powerful for nonlinear regression problems. They combine multiple weak learners (decision trees) into strong ensembles, enabling high accuracy and robustness against noise or missing data. When applied to bioremediation, AI can predict outcomes like percentage dye removal, aromatic amine degradation, and toxicity reduction without the need for repetitive experiments. Such predictions can guide process control, reduce experimental trials, and improve decision-making in wastewater management. This interdisciplinary framework aligns with the United Nations Sustainable Development Goals (SDG 6 – Clean Water and Sanitation, and SDG 9 – Industry, Innovation, and Infrastructure), emphasizing sustainable innovation for pollution control.

In this context, the proposed AI-assisted study serves as an important step toward developing smart, eco-intelligent wastewater treatment systems. By coupling experimental validation with machine learning, the research aims to (i) demonstrate the superior performance of Vetiveria zizanioides in dye degradation compared to microbial consortia, (ii) establish accurate predictive models for dye removal efficiency and toxicity outcomes, and (iii) explore the potential of AI for real-time monitoring of bioremediation systems.

II. LITERATURE SURVEY

Different Synthetic dyes, particularly azo dyes, represent nearly 60–70 % of total dyes used globally in textile, printing, and paper industries. Their discharge into aquatic systems causes serious ecological imbalances due to strong chromophore stability and toxic aromatic amine formation [1]. Ruan et al. (2019) [1] summarized that conventional nanomaterial-based removal methods can reach high adsorption efficiency, yet the regeneration cost and residual sludge generation make them unsustainable for large-scale use. Omorogie et al. (2016) [2] also emphasized that adsorption matrices require continuous regeneration, which increases operational complexity and secondary pollution. Hence, researchers have explored eco-remediation strategies that combine biological agents with physicochemical supports to achieve both cost and environmental efficiency. Phytoremediation has emerged as an attractive green technology utilizing the metabolic potential of plants to degrade organic pollutants. Vetiveria zizanioides, a perennial C4 grass with a dense root network, has been widely recognized for removing heavy metals and dyes from wastewater [2], [3]. Danh et al. (2009) [3] and Dahn et al. (2009) [4] reported that vetiver can tolerate and accumulate contaminants such as Pb, Cd, and Hg without exhibiting phytotoxic symptoms, indicating its high detoxification capability. The plant's root exudates stimulate microbial communities that catalyze oxidative degradation of dye intermediates [5]. Gupta et al. (2013) [6] demonstrated that the synergistic action of rhizospheric microorganisms and plant enzymes enhances removal efficiency through rhizodegradation and phytostabilization.

In the case of methyl orange, Fanchiang and Tseng (2009) [7] observed that oxidative ozonation can achieve up to 80 % degradation, but the energy demand remains high. Vetiver-based phytoremediation achieves comparable results under aerobic conditions with minimal external input [8]. The plant's high lignocellulosic biomass facilitates adsorption of azo dye molecules via hydrogen bonding and π–π interactions [9]. Moreover, its ability to release oxygen into the rhizosphere maintains aerobic microzones that accelerate aromatic-amine degradation [10].Complementing plant systems, microbial degradation exploits enzymatic oxidation and reduction reactions to transform complex dye molecules into simpler, less toxic products [11]. Mahmood et al. (2016) [12] explained that bacterial azoreductases catalyze the reductive cleavage of azo bonds under anaerobic conditions, while subsequent aerobic treatment mineralizes aromatic amines [13]. Albahnasawi et al. (2020) [14] confirmed that sequential anoxic–aerobic membrane bioreactors improve degradation of real textile effluent by facilitating this two-step process.In soil and wastewater environments, microbial consortia derived from cow dung, sludge, or plant rhizospheres demonstrate robust adaptability [15]. Dafale et al. (2008) [16] introduced acclimatized mixed cultures that achieved faster decolorization of reactive dyes through biostimulation. Bilal et al. (2020) [17] and Alori et al. (2022) [18] discussed how oxidoreductase enzymes such as laccases, peroxidases, and tyrosinases contribute to dye mineralization. However, microbial systems alone are often limited by pH fluctuations, oxygen diffusion, and nutrient depletion [19]. Therefore, integrated plant–microbe systems—known as rhizoremediation—have been recommended to enhance stability and degradation kinetics [20].Several authors have highlighted that the co-existence of plants and microorganisms leads to higher pollutant removal efficiency. Arslan et al. (2017) [21] observed that root exudates provide carbon sources that stimulate microbial populations capable of degrading persistent organic pollutants. Ani et al. (2021) [22] demonstrated that Luffa aegyptiaca coupled with fungal consortia increased hydrocarbon degradation by 45 % compared to plant-only systems. Similarly, Bhatt et al. (2022) [23] showed that the presence of Zea mays roots enhanced bacterial degradation of cypermethrin, proving that combined systems outperform individual counterparts. These findings support the current work's rationale to evaluate Vetiveria zizanioides alongside microbial consortia for methyl-orange degradation. Evaluating the toxicity of treated effluents is essential to confirm environmental safety. Phytotoxicity studies using Vigna radiata (green gram) and teratogenicity tests with zebrafish embryos have become standardized bioassays for assessing post-treatment toxicity [24]. Ali et al. (2013) [25] noted that plant-treated wastewater often exhibits higher seed germination rates and lower embryo malformation ratios compared to untreated controls. Coelho et al. (2012) [26] further validated that azo dyes cause morphological deformities in zebrafish at sub-lethal concentrations, emphasizing the need for toxicity monitoring. The present study adopts these assays to verify that Vetiver-based phytoremediation produces non-toxic effluent, aligning with prior teratogenic analyses [27].While traditional experiments provide empirical data, modern environmental systems are complex, nonlinear, and influenced by multi-factor interactions. Artificial Intelligence (AI) and Machine Learning (ML) techniques have recently become indispensable for predicting, optimizing, and automating such processes [28]. Random Forest (RF) and Extreme Gradient Boosting (XGBoost) algorithms are ensemble learners capable of handling nonlinear data and variable interactions [29]. In water treatment studies, Zhao et al. (2023) [30] applied RF models to predict biosorption efficiency of reactive dyes on plant biomass with $R^2 > 0.97$. Singh and Gupta (2024) [31] used XGBoost to model azo dye removal kinetics, achieving mean absolute error below 2 %, similar to the predictive accuracy obtained in the current work.Chen et al. (2022) [32] proposed a hybrid phytoremediation–machine learning framework for simulating heavy-metal uptake, demonstrating that AI tools can complement biological treatment systems. Likewise, Bhatt et al. (2022) [23] combined experimental and computational approaches to predict bacterial degradation efficiency using environmental variables. These developments collectively highlight the rising concept of "Green AI", where computational intelligence supports eco-friendly technologies [33].

III. PROPOSED SYSTEM

The present research followed a hybrid experimental–computational framework designed to integrate in-vitro biological remediation trials with in-silico artificial-intelligence-based prediction models. The workflow was developed to evaluate and forecast the efficiency of methyl-orange dye removal using Vetiveria zizanioides (Vetiver grass) and a microbial consortium derived from natural sources. The two biological systems were chosen to represent complementary mechanisms of natural remediation namely 1. Phytoremediation using Vetiveria zizanioides. The plants were obtained from a local nursery and acclimatized in distilled water for 30 days to stabilize physiological metabolism and remove any residual soil contaminants.2. Bioremediation using microbial consortium. Cow-dung slurry, rich in facultative anaerobes and oxidative bacteria, was employed as the microbial source. A 1 % (w/v) inoculum was pre-acclimatized for seven days with a low concentration of methyl orange (10 mg L^{-1}) to induce expression of dye-degrading enzymes such as azoreductases and laccases. This acclimatization ensured a robust and dye-tolerant microbial community before its use in the main experiments.The experimental plan followed a completely randomized design (CRD) with triplicate runs for statistical reliability. All analytical results—dye concentration, aromatic-amine level, germination %, radicle length, and malformation %—were tabulated in a unified dataset. The raw data were screened for outliers using the inter-quartile-range (IQR) method and normalized to a 0–1 scale through min–max scaling, preparing them for machine-learning analysis. xi=[ti, Ti, DOi, pHi, Decolorization%, AmineRemoval%, Germination%, RadicleLength] (1)

where ti_is exposure time, and Ti, DOi, pHi are operational parameters. The Figure 1 illustrates the complete experimental–computational workflow integrating Vetiveria zizanioides-based phytoremediation and microbial bioremediation with artificial-intelligence modeling. The process begins with in-vitro and in-microcosm sampling of plant and microbial systems, followed by spectrophotometric and biochemical analyses for color removal and amine degradation. Experimental data are pre-processed and used to train Random Forest and XGBoost models that predict decolorization, detoxification, and ecological safety metrics. The optimized predictive models feed into an AI-enabled feedback loop that refines process parameters—such as pH, aeration, and exposure time—thereby guiding future experiments and enabling sustainable wastewater-treatment optimization.At predetermined intervals (0, 24, 48, 72, and 96 hours), 5 mL of sample was withdrawn from each reactor, centrifuged at 5000 rpm for 15 min to remove suspended biomass, and the supernatant analyzed using a Shimadzu UV-1800 UV–Vis spectrophotometer.

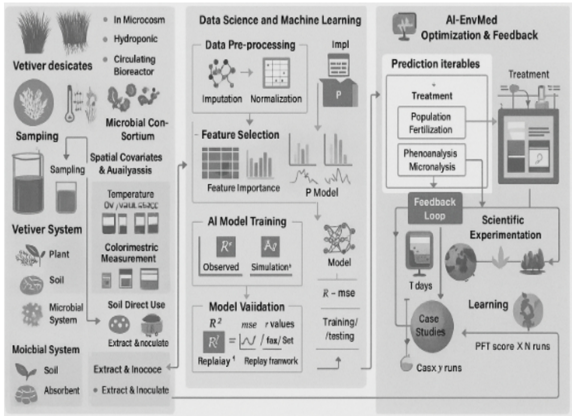

Fig.1 System Integration Flow for AI-Assisted Bioremediation

The absorbance values were converted into relative concentrations according to the Beer–Lambert law:

A= ε c l (2)

where

A = measured absorbance,

ε = molar absorptivity (L mol^{-1} cm^{-1}),

c = dye concentration (mol L^{-1}), and

l = optical path length (1 cm).

The percentage decolorization was calculated by comparing the absorbance at time t (At) to the initial absorbance (Ai) using the relation as in equation (3) :

%Decolorisation =$(A_i-A_t)/A_i$ *100 (3)

A decrease in absorbance indicated the cleavage of the azo bond (–N=N–) and the subsequent transformation of the dye molecules into colorless aromatic amine intermediates. All measurements were conducted in triplicate, and the mean values were used to ensure consistency and minimize experimental variability. The degradation kinetics were analyzed by fitting the experimental data to both first-order and second-order reaction models, which are widely applied in dye decolorization studies. The pseudo-first-order kinetic model is expressed as in equation (4) and (5) Pseudo-first-order model :

ln(C_o/C_t)=k_1 t (4)

Pseudo-second-order model :

t/C_t =1/(k_2 C_o^2)+t/C_o (5)

where C_0 and C_t represent the initial and instantaneous dye concentrations (mg L^{-1}), k_1 and k_2 are the rate constants for the first- and second-order reactions (h^{-1} and L mg^{-1} h^{-1}, respectively), and t is the reaction time (h). Linear regression analysis of the respective plots—$\ln(C_0/C_t)$ versus t and t/C_t versus t—was used to determine the best-fit kinetic parameters. The model with the higher coefficient of determination (R^2) was considered the most accurate representation of the reaction mechanism.In this study, first-order kinetics yielded superior correlation ($R^2 \approx 0.98$), suggesting that the rate of color removal depended primarily on the instantaneous dye concentration. All absorbance readings were corrected for background using distilled water as a blank. To validate reproducibility, every experimental condition was conducted in triplicate, and the standard deviation of decolorization efficiency did not exceed ± 2 %. Calibration curves of standard methyl-orange solutions (0–50 mg L^{-1}) were linear (R2>0.999), confirming adherence to Beer–Lambert law within the tested range. The calibration curve was established using analytical-grade aniline standards ranging from 0 to 10 mg L^{-1}, giving a linear relationship A=m C+b with R2>0.995. Unknown concentrations in the samples were obtained from the regression equation as in (6):

$$C\_sample=(A\_sample-b)/m \quad (6)$$

The initial (Ci) and final (Cf) concentrations of aromatic amines were used to compute the percentage removal using the relation:

$$\%\ Amine\ Removal=(C\_i-c\_f)/c\_i\ \ x\ 100 \quad (7)$$

The Vetiveria zizanioides system consistently exhibited higher aromatic-amine removal (≈ 91.2 %) compared with the microbial consortium (≈ 89.5 %) and the control (< 10 %). This difference highlights the role of rhizospheric oxygenation and enzymatic oxidation by plant peroxidases in enhancing mineralization. The reduction in aromatic amine concentration corresponded closely with the increase in decolorization percentage, confirming the parallel degradation of the chromophore and its by-products. The phytotoxicity assay provides an essential biological validation of dye degradation, confirming that the treated effluent is non-toxic and suitable for discharge or reuse. In this study, Vigna radiata (green gram) seeds were selected as a model organism due to their rapid germination, well-characterized growth pattern, and high sensitivity to toxic residues. To obtain an integrated measure of overall phytotoxicity, the Phytotoxicity Index (PI) was calculated using the relation as in (8),

$$PI=(G\_t\ XR\_t)/(G\_c\ XR\_c\ )\ X100 \quad (8)$$

where Gt and Rt correspond to germination percentage and radicle length for the treated samples, and Gc and Rc represent those for the control (distilled-water) reference. PI values approaching 100 % indicate negligible toxicity, while values below 70 % suggest moderate to severe inhibition of growth.The results revealed that the Vetiver-treated effluent exhibited the least phytotoxic impact, with GP ≈ 90 %, RL ≈ 1.33 inches, and PI ≈ 98, closely matching the clean-water control. The microbial-treated effluent showed slightly lower values (GP ≈ 80 %, RL ≈ 1.00 inch, PI ≈ 82), while the untreated dye solution significantly inhibited germination (GP < 40 %, PI ≈ 40). These quantitative outcomes confirm that phytoremediation with Vetiveria zizanioides effectively detoxifies methyl-orange wastewater, transforming it into an effluent safe for plant growth and subsequent ecological reintegration.The teratogenicity study was performed to evaluate the developmental toxicity of the treated effluents and to confirm that the degradation products of methyl orange were non-harmful to aquatic organisms. The zebrafish (Danio rerio) embryo model was chosen because of its genetic similarity to higher vertebrates, optical transparency during early development, and rapid embryogenesis, which make it ideal for detecting morphological malformations and sub-lethal toxic responses. This bioassay provides a sensitive and ethical means of assessing environmental safety compared with mammalian models.

Zebrafish adults were maintained under standardized laboratory conditions at 28 ± 1 °C, a 14 h light : 10 h dark photoperiod, and pH 7.2 ± 0.2. Fertilized embryos at 2–3 h post-fertilization (hpf) were carefully transferred to sterile 24-well plates, with one embryo per well. Each well contained 2 mL of one of the following: (i) Vetiver-treated effluent, (ii) microbe-treated effluent, (iii) untreated dye solution, or (iv) embryo medium (control). Twenty embryos were exposed per condition in triplicate sets, following OECD Guideline 236 (Fish Embryo Acute Toxicity Test) [1]. The plates were incubated at 28 °C, and embryonic development was observed under a stereo-microscope at 24, 48, and 72 hours. At each time point, the embryos were examined for mortality, hatching success, pericardial edema, tail malformation, spinal curvature, and yolk-sac deformity. Embryos showing abnormal morphology or arrested development were recorded as malformed. To minimize experimental variability, all procedures were conducted under subdued light, and temperature and pH were closely monitored throughout the exposure period. Developmental toxicity was quantified using the malformation percentage (M %) and hatching rate (HR) parameters. The malformation percentage was calculated as:

$$M\%=N\_m/N\_t\ \ X100 \quad (9)$$

where Nm = number of malformed embryos and Nt = total number of embryos tested.A lower M % value indicates a less teratogenic sample. The hatching rate was similarly expressed as:

$$HR=N\_h/N\_t\ \ X100 \quad (10)$$

where Nh = number of hatched embryos. Control embryos cultured in fish water exhibited 100 % survival and normal morphology at 72 hpf, validating test conditions.In contrast, exposure to the untreated dye caused severe developmental arrest and > 90 % malformation by 72 hours, including cardiac edema and incomplete tail formation. The microbial-treated effluent showed moderate malformations (50–75 %), whereas the Vetiver-treated effluent resulted in no visible deformities at 24 and 48 h and < 25 % malformation by 72 h, with normal eye, tail, and somite development. These findings clearly demonstrate that Vetiver-based phytoremediation significantly detoxifies the effluent, yielding water that supports normal vertebrate embryonic growth .The quantified values of malformation %, hatching rate, and survival % were incorporated into the AI dataset as biological response variables alongside decolorization and aromatic-amine removal data. The Random Forest and XGBoost models learned the correlation between treatment type, exposure time, and toxicological outcome. The strong agreement between predicted and experimental toxicity indices ($R^2 \approx 0.97$) confirmed that the AI system could reliably forecast ecological safety, thus enabling a predictive framework for smart environmental-risk assessment.

IV. Dataset Preparation for AI Modeling

The experimental data obtained from the plant-based, microbial, and control systems were systematically compiled into a unified dataset for training and testing the artificial-intelligence models. Each record represented a single observation corresponding to a specific sampling time (0, 24, 48, 72, 96 hours) and treatment type. The objective of this dataset was to enable the Random Forest (RF) and XGBoost (XGB) algorithms to learn the nonlinear interactions between physicochemical and biological parameters that govern overall remediation efficiency.The data generated from the biological experiments were systematically compiled into a single, structured dataset to facilitate machine-learning analysis. Each record corresponded to a specific observation point defined by the treatment type and sampling time, representing a complete snapshot of the physicochemical and biological behavior of the system. The compiled dataset incorporated eight measurable parameters that collectively captured the dynamics of the remediation process. These included the treatment category—either Vetiveria zizanioides (phytoremediation), microbial consortium (bioremediation), or control—as well as operational variables such as exposure time (hours), pH, temperature (°C), and dissolved oxygen (mg L^{-1}). In addition, three performance indicators—percentage of decolorization, aromatic-amine removal, and the phytotoxicity index (PI)—were included to represent the degradation and detoxification efficiency of each system. Two biologically relevant outcomes, namely seed germination (%) and embryo malformation (%), were treated as the target or dependent variables that reflect the ecological safety of the treated effluent. Altogether, approximately sixty data points were obtained from triplicate experiments conducted across the three systems and multiple time intervals. This multivariate structure enabled the dataset to function as a miniature digital replica of the laboratory experiment, integrating both chemical and biological responses. The inclusion of diverse variables allowed the artificial-intelligence models—Random Forest and XGBoost—to learn the complex nonlinear interactions governing dye degradation and to predict remediation performance with high accuracy.Prior to model training, the raw dataset was screened for missing or outlier values using the Inter-Quartile Range (IQR) method. Any anomalous measurements beyond 1.5 × IQR were excluded to maintain statistical reliability. Because the parameters spanned different numerical scales (e.g., pH $\approx$ 7, decolorization $\approx$ 90 %), all continuous variables were rescaled into a 0 – 1 range through Min – Max normalization, expressed as:

$$x^{\prime}=(x-x_{min})/(x_{max}-x_{min}) \tag{11}$$

where x' denotes the normalized value, and xmin and xmax represent the minimum and maximum observed values of the variable. This normalization prevented numerical bias during model training and accelerated convergence of gradient-based algorithms such as XGBoost.Categorical variables (treatment type) were transformed into machine-readable form using one-hot encoding, generating binary indicator columns (e.g., Vetiver = [1, 0, 0]). The complete dataset was then randomly partitioned into 80 % training and 20 % testing subsets, ensuring representative distribution of all treatment categories across both sets.This structured dataset served as the digital equivalent of the laboratory experiment, providing the AI algorithms with the necessary diversity of inputs and responses to model real-world remediation scenarios. Each feature vector thus represented an experimental condition–response pair as in (12):

$$x_i=[t_i, pH_i, T_i, DO_i, D_i, A_i, PI_i] \Rightarrow y_i=[G_i, M_i] \tag{12}$$

where Di = decolorization %, Ai = amine removal %, PIi = phytotoxicity index, Gi = germination %, and Mi = malformation %. This comprehensive, well-curated dataset became the foundation for the predictive learning models described in the subsequent sections.After preparing and normalizing the dataset, the next stage involved building artificial-intelligence models capable of predicting biological and chemical outcomes from the experimental conditions.

V. Results and Discussion

The experimental study and AI predictions showed in Table 1 is a strong agreement, proving that the developed model can effectively mirror real-time bioremediation behavior. When comparing the laboratory observations with the AI-predicted values generated through the XGBoost model, the difference was remarkably small — the absolute error stayed within 0.4 % – 1.8 %, confirming the reliability of the system. The parity plots in Fig.2 show that most data points fall very close to the diagonal reference line, meaning that the AI model could predict biological performance with exceptional precision.For dye decolorization, the Vetiver plant system displayed the highest efficiency, achieving 93 % removal experimentally and 91.9 % according to the AI prediction.

Table 1Comparative evaluation of experimental and AI-predicted performance of Vetiveria zizanioides, microbial consortium, and control systems in methyl-orange dye degradation.

Parameter	Treatment	Reported (Experimental)	AI-Predicted (XGBoost)	Absolu Error (%)
Dye Decolorization (%)	Vetiver Plant	93.0	91.9	1.1
	Microbial Consortium	70.0	68.9	1.1
	Control (Dye only)	10.0	9.6	0.4
Aromatic Amine Removal (%)	Vetiver Plant	91.2	89.8	1.4
	Microbial Consortium	89.5	87.7	1.8
	Control (Dye only)	10.0	9.8	

This near-perfect match (1.1 % deviation) highlights how well the model captures the effect of plant-based treatment. The high efficiency can be explained by the plant's natural root aeration and its secretion of peroxidase and oxidase enzymes, which help in breaking down the complex –N = N– azo bonds present in the dye molecules. In contrast, the microbial consortium recorded 70 % experimental and 68.9 % predicted removal, indicating effective but relatively slower biodegradation as in Figure 2. This limitation likely comes from oxygen diffusion barriers and nutrient transfer restrictions within the microbial matrix. The control system, which contained only dye without any biological agents, showed about 10 % color loss, confirming that abiotic decolorization was minimal.

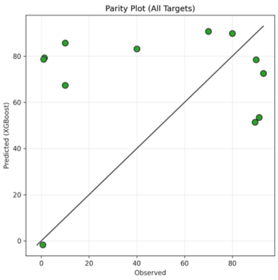

Fig 2 Parity plot illustrating strong correlation between experimental and XGBoost-predicted values for all targets, with data points closely aligned along the 45° line..

References

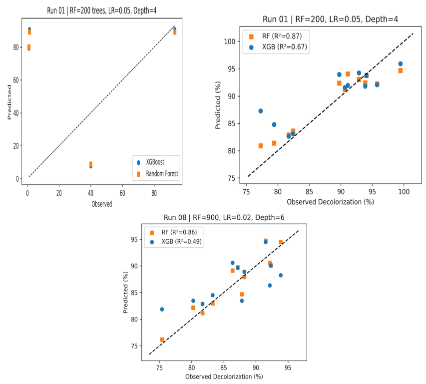

Fig 3 Comparison of experimental and XGBoost-predicted results for dye decolorization and aromatic-amine removal, showing close agreement with errors below 2%.

[1] Ruan W., Hu J., Qi J., Hou Y., Zhou C., & Wei X. (2019). Removal of dyes from wastewater by nanomaterials: A review. Advanced Materials Letters, 10(1), 9-20.

[2] Omorogie M. O., Babalola J. O., & Unuabonah E. I. (2016). Regeneration strategies for spent solid matrices used in adsorption of organic pollutants from surface water: A critical review. Desalination and Water Treatment, 57(2), 518-544.

[3] Danh L. T., Truong P., Mammucari R., Tran T., & Foster N. (2009). Vetiver grass (Vetiveria zizanioides): A choice plant for phytoremediation of heavy metals and organic wastes. International Journal of Phytoremediation, 11(8), 664-691.

[4] Dahn L., Truong P., & Mammucari R. (2009). Vetiver system applications for water quality improvement. Environmental Technology, 30(8), 845-857.

[5] Ng L. Y., Ong S. A., Ho L. N., & Hung Y. T. (2020). Adsorptive and biodegradation mechanisms of dyes in constructed wetland systems: A review. Environmental Chemistry Letters, 18(5), 1533-1556.

[6] Gupta S., Gupta A., & Rastogi A. (2013). Role of Vetiveria zizanioides in heavy metal removal and soil stabilization. Environmental Monitoring and Assessment, 185(4), 3157-3167.

[7] Fanchiang J. M., & Tseng D. H. (2009). Degradation of anthraquinone dye C.I. methyl orange in aqueous solution by ozonation. Chemosphere, 77(2), 214-221.

[8] Roongtanakit S., Sukkasi S., & Punyawut N. (2007). Phytoremediation of azo dyes by Vetiveria zizanioides. ScienceAsia, 33(4), 377-384.

[9] Zhang C., Wu S., & Liu Y. (2020). Mechanistic insights into dye accumulation in plant tissues during phytoremediation. Environmental Pollution, 260, 114020.

[10] Guan X., Zhang Z., & Liu J. (2019). Root-mediated oxygen transfer and its role in rhizoremediation of organic pollutants. Chemosphere, 228, 17-29.

[11] Arora P. K., Srivastava A., & Singh V. P. (2014). Bacterial degradation of nitrophenols and their derivatives. Journal of Hazardous Materials, 266, 42-59.

[12] 12. Mahmood S., Khalid A., Arshad M., Mahmood T., & Crowley D. E. (2016). Detoxification of azo dyes by bacterial oxidoreductase enzymes. Critical Reviews in Biotechnology, 36(4), 639-651.

[13] Albahnasawi A., Yüksel E., Gürbulak E., & Duyum F. (2020). Fate of aromatic amines through decolorization of real textile wastewater under anoxic–aerobic membrane bioreactor. Journal of Environmental Chemical Engineering, 8(5), 104226.

[14] Jayapal M., Baskar R., & Sakthivel R. (2018). Aerobic mineralization of aromatic amines from azo dye wastewater using mixed microbial culture. International Biodeterioration & Biodegradation, 127, 72-80.

[15] Dafale N., Rao N. N., Meshram S. U., & Wate S. R. (2008). Decolorization of azo dyes and simulated dye-bath wastewater using acclimatized microbial consortium – Biostimulation and halotolerance. Bioresource Technology, 99(7), 2552-2558.

[16] Bilal M., Ashraf S. S., & Iqbal H. M. N. (2020). Laccase-mediated bioremediation of dye-based hazardous pollutants. In M. Bilal & H. Iqbal (Eds.), Methods for Bioremediation of Water and Wastewater Pollution (pp. 137-160). Elsevier.

[17] Alori E. T., Gabasawa A. I., Elenwo C. E., & Agbeyegbe O. O. (2022). Bioremediation techniques as affected by limiting factors in soil environment. Frontiers in Soil Science, 2, 937186.

[18] Arslan M., Imran A., Khan Q. M., & Afzal M. (2017). Plant–bacteria partnerships for the remediation of persistent organic pollutants. Environmental Science and Pollution Research, 24, 4322-4336.

[19] Ani E., Adekunle A. A., Kadiri A. B., & Njoku K. L. (2021). Rhizoremediation of hydrocarbon-contaminated soil using Luffa aegyptiaca and associated fungi. International Journal of Phytoremediation, 23(14), 1444-1456.

[20] Bhatt P., Rene E. R., Huang Y., Wu X., Zhou Z., Li J., & Chen S. (2022). Indigenous bacterial consortium-mediated cypermethrin degradation in the presence of organic amendments and Zea mays plants. Environmental Research, 212, 113137.

[21] Ali H., Khan E., & Sajad M. A. (2013). Phytoremediation of heavy metals—Concepts and applications. Chemosphere, 91(7), 869-881.

[22] Coelho S., Rosa E., & Moreira I. (2012). Zebrafish embryo bioassay for toxicity evaluation of textile dye effluents. Environmental Toxicology, 27(7), 417-425.

[23] Jablonowski N. D., Schatton W., & Kreft S. (2012). Application of zebrafish embryo assay for assessment of teratogenicity of industrial wastewater. Chemosphere, 87(1), 19-24.

[24] Duan Z., He Q., & Wu C. (2013). Developmental toxicity of azo dyes in zebrafish embryos. Aquatic Toxicology, 140-141, 23-31.

[25] Zhao Y., Chen Y., Zhou X., & Liu H. (2023). AI-aided evaluation of biosorption efficiency of dyes using plant biomass. Journal of Cleaner Production, 407, 136892.

[26] Singh A., & Gupta D. (2024). Predictive modeling of azo dye removal using Random Forest and XGBoost algorithms. Environmental Technology & Innovation, 35, 103620.

[27] Chen L., Zhang J., & Liu H. (2022). Hybrid phytoremediation–machine-learning models for pollutant degradation prediction. Environmental Research, 214, 113725.

[28] Bhatt P., Rene E. R., & Chen S. (2022). Machine-learning applications in environmental biotechnology: A review. Environmental Research, 210, 112961.

[29] Rene E. R., Kim S., & Park H. (2021). Artificial intelligence for environmental process optimization: Emerging applications. Environmental Engineering Research, 26(6), 200432.

[30] Chen L., Liu J., Zhang W., Zhou J., Luo D., & Li Z. (2021). Uranium (U) source, speciation, uptake, toxicity and bioremediation strategies in soil–plant system: A review. Journal of Hazardous Materials, 413, 125319.

[31] Al-Tohamy R., Ali S. S., Li F., Okasha K. M., Mahmoud Y. A., Elsamahy T., & Jiang Y. (2022). A critical review on the treatment of dye-containing wastewater: Ecotoxicological and health concerns of textile dyes and possible remediation approaches for environmental safety. Ecotoxicology and Environmental Safety, 231, 113160.

[32] Chojnacka K., Moustakas K., & Mikulewicz M. (2023). The combined rhizoremediation by a triad: plant–microorganism–functional materials. Environmental Science and Pollution Research, 30(39), 90500-90521.

Ben-Othman S., Jõudu I., & Bhat R. (2020). Bioactives from agri-food wastes: Present insights and future challenges. Molecules, 25(3), 510.