Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Artificial Intelligence-Assisted Prediction and Optimization of Methylene Blue Adsorption onto Albizia lebbeck Bio-adsorbent Using Ensemble Machine Learning

**Sayali Kawalkar**
Ramdeobaba University
**Swarooplaxmi Mudliar**
mudliarsl@rknec.edu
Ramdeobaba University
**Dhanashree Kuthe**
Ramdeobaba University
**Sai Shashank Mudliar**
Purdue University Fort Wayne
**Ashish Kamble**
Ramdeobaba University

**Article**

**Keywords:** Methylene Blue; Albizia lebbeck; Bio-adsorption; Random Forest; eXtreme Gradient Boosting (XGBoost); Ensemble Machine Learning; Wastewater Remediation

**Posted Date:** November 3rd, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-10956942/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

**Additional Declarations:** No competing interests reported.

# Artificial Intelligence-Assisted Prediction and Optimization of Methylene Blue Adsorption onto *Albizia lebbeck* Bio-adsorbentUsing Ensemble Machine Learning

SayaliKawalkar[1],SwarooplaxmiMudliar[1,*],DhanashreeKuthe[2],SaiShashank Mudliar[3], AshishKamble[4]

[1]Department of Chemistry, Ramdeobaba University, Nagpur, Maharashtra, India 440013

[2]Department of Computer Science, Ramdeobaba University, Nagpur, Maharashtra,India440013

[3]Department of Computer Science, Purdue University Fort Wayne, Fort Wayne, IN, USA

[4]Department of Computer Science, Ramdeobaba University, Nagpur ,Maharashtra,India440013

*CorrespondingAuthor:mudliarsl@rknec.edu

**Abstract**

This study investigates the adsorption of methylene blue (MB)from aqueous solution using a bio-adsorbent prepared from *Albizialebbeck* pods and evaluates the predictive performance of ensemble machine learning models for modeling the adsorption process. The bio-adsorbent was characterized using pHPZC, Fourier Transform Infrared spectroscopy (FTIR), X-ray Diffraction (XRD),and Scanning Electron Microscopy with Energy Dispersive Spectroscopy (SEM–EDS). Characterization revealed a porous, carbon-rich, predominantly amorphous structure with abundant oxygen-containing functional groups favourable for dye adsorption. The effects of solution pH, contact time, adsorbent dosage, and initial dye concentration were systematically investigated through batch adsorption experiments. Equilibrium and kinetic data were analysed using Langmuir, Freundlich, and Temkin isotherm models, as well as pseudo-first-order, pseudo-second-order, and intraparticle diffusion kinetic models. The Langmuir isotherm provided the best equilibrium fit, yielding a maximum monolayer adsorptioncapacityof424.24mgg$^{-1}$, while the pseudo-second-order model best described the adsorption kinetics. Furthermore, ensemble machine learning models—Random Forest (RF) and eXtreme Gradient Boosting(XGBoost)—were developed and optimized to predict adsorption performance. The optimized RF model achieved an $R^2$value of 0.98876 for MB removal efficiency, while XGBoost attained an $R^2$of0.99953 for adsorption capacity prediction. These findings demonstrate that *Albizia lebbeck* bio-adsorbent is a highly efficient, eco-friendly material for MB removal, and that ensemble learning provides a robust and accurate framework for predicting and optimizing adsorption processes in wastewater treatment.

**Keywords:** Methylene Blue; *Albizia lebbeck*; Bio-adsorption; Random Forest; eXtreme Gradient Boosting (XGBoost); Ensemble Machine Learning; Wastewater Remediation

## 1. Introduction

Water contamination by synthetic dyes represents one of the most pressing environmental challenges of the 21st century, threatening aquatic ecosystems and human health on a global scale (1).Water and the environment constitute two fundamental life-sustaining resources,andtheirdegradationduringindustrial processes—particularly textile manufacturing—carries severe ecological consequences. As global population growth drives increasing per capita consumption of textile products, the demandfordyesandassociatedchemicalscontin-ues to escalate at an alarming rate (2). The textiledyeingindustry, especiallyindeveloping nations, consumes millions oftonnes of freshwater annually while discharging vast quantities of contaminated effluent containing residual dyestuffs, hazardous chemicals, and gaseous emissions. Moreover,chemicalspeciesmaybe released during the processing, service life, or end-of-life disposal of finished dyed products, compounding the environmental burden (3).

Among the numerous synthetic dyes used in industry, methylene blue (MB) is of particular concern owing to its widespread use, environmental persistence, and documented toxicity. Methyleneblueisanon-biodegradablecationic dyebelongingtothethiazinefamily,characterized by a fused-ring structure containing nitrogen and sulfur heteroatoms and a conjugated$\pi$-electron system, with the molecular formula $C_{16}H_{18}N_3SCl$ (4). In the MB molecule, the chromophore arises from conjugated interactionsbetweenthenitrogenandsulfuratoms in the central heterocyclic ring, while the auxochromic propertiesare conferred by nitrogen-containinggroupsonthebenzeneringthatpos-sess lone electron pairs (5).Exposure to MB can cause ocular injuries resembling chemical burns—potentiallyleadingtopermanentvision loss in bothhumans and animals—as well as respiratorydistress,dermalsensitization,cyanosis, diaphoresis, dizziness, and emesis (6; 7).

Various treatment technologies have been employed for the removal of dye contaminants fromwastewater,broadlyclassifiedintobiological methods (biodegradation), chemical methods (coagulation, flocculation, chemical precipitation, and oxidation), and physical methods (filtration, adsorption, and membrane separation) (8;9). Among these, adsorption has emergedasthemost attractiveapproachowing toseveralinherent advantages: effectiveness at very low concentrations, applicability in both batch and continuous process configurations, operationalsimplicity,adsorbentregenerability and reusability, and comparatively low investment costs (10).

Concurrently,extensiveresearcheffortshave beendevotedtodevelopingpredictivemod-els for water quality assessment to enhance management strategies and early warning systems (11). However, a persistent challenge in water treatment modelling lies in the inherently nonlinear relationships and fluctuations characteristic of water-related data (12). Artificial intelligence (AI) has emerged as a transformativetoolinthisdomain,offeringease of implementation, cost-effectiveness, the abil-ity to handle large and complex datasets, and the capacity to model highly nonlinear phenomenathatcannotbeadequatelycaptured by conventional empirical equations (13). Numerous recent studies have demonstrated the increasing adoption of AI approaches for water treatment modelling, employing algorithms suchasartificialneuralnetworks(ANNs),multilinearregression(MLR), radial basisfunction (RBF) networks, and support vector machines

(SVMs) for water quality prediction and monitoring (14; 15; 16).

*Albizia lebbeck*is a widely distributed deciduous tree found in tropical and subtropical regions, belonging to the family Fabaceae and producing seed pods measuring 10–20 cm in length. Its pod biomass constitutes a high-volume agricultural waste that poses disposal challenges; however, its composition—36.4% cellulose, 18.9% hemicellulose, 13.6% lignin, and 83.1% volatile matter—makes it an excellentprecursorforbio-adsorbentprepara-tion (17).

In the present study, a green bio-adsorbent wassynthesizedfrom*Albizialebbeck*podsand comprehensivelycharacterizedfortheefficient removal of methylene blue from aqueous solution. The physicochemical properties of the prepared bio-adsorbent were evaluated using $pH_{PZC}$, FTIR, XRD, and SEM–EDS analyses. The influence of solution pH, contact time, adsorbent dosage, and initial dye concentration on adsorption behaviour was systematically investigated through batch experiments, and the results were interpreted using equilibrium isotherm and kinetic models. Additionally, ensemble machine learning models—specifically Random Forest and eXtreme Gradient Boosting—were developedand validated for predicting methylene blue removalefficiency and adsorption capacity from experimental data.

### ResearchObjectives

The specific objectives of this study are as follows:

(i) To synthesize and characterize a bio-adsorbent derived from*Albizia lebbeck* pods using $pH_{PZC}$, FTIR, XRD, and SEM–EDSformethyleneblueadsorption applications.

(ii) Toinvestigatetheinfluenceofsolution pH,contacttime,adsorbentdosage,and initialdyeconcentrationontheadsorption efficiency.

(iii) To elucidate the adsorption mechanism throughequilibriumisothermandkinetic modelling.

(iv) Todevelop,optimize,andcomparatively evaluate Random Forest and XGBoost modelsforaccuratepredictionofmethyleneblueremovalefficiency andadsorption capacity.

## 2. Methodology

### Materials,BotanicalAuthentication, and Bio-adsorbent Preparation

Mature seed pods of *Albizia lebbeck* (L.) Benth. (Family: Fabaceae) were collected from natural tree habitats in Nagpur, Maharashtra, India. The plant material was botanically identified by Dr. Nitin Dongarwar, Professor, Department of Botany, Rashtrasant Tukadoji Maharaj Nagpur University, Nagpur, Maharashtra, India. The identification was done using the typical botanical and morphological features of the plant such as pod and others diagnostic features. The identity of the plant has been confirmed as *A. lebbeck* . A voucher specimen was deposited, giving voucher specimen number 2026/559 (Figure1b).Figure1.

The harvested pods were subjected to a rigorouscleaningprocessusingdistilledwater to eliminate adhering surface particulates, followed by shade-drying and subsequent sun-drying for 7 days under controlled ambient humidity.Thedriedpodswerethenplaced inan incubation oven at 105°C for 24 hours to ensurecompletemoistureelimination.Thedesiccatedmaterialwasmechanicallygroundusing alaboratory mill and mortar, followed by siev-





(a) Treecanopy&pods(b)Herbariumsheet

Figure 1: Botanical identification and voucher authentication of *Albizia lebbeck*(L.) Benth.: (a) mature seed pods hanging from tree canopy in natural habitat; and (b) authenticated herbarium voucher specimen (Sheet No. 2026/559, Department of Botany, RTMNU Nagpur University).

ing to achieve a uniform particle size fraction of<425 µm (Figure2). The prepared bio-adsorbent was stored in airtight containers to preserve structural integrity and prevent atmospheric moisture absorption until required for characterization and adsorption experiments.



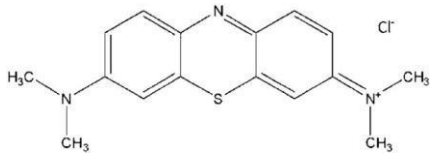
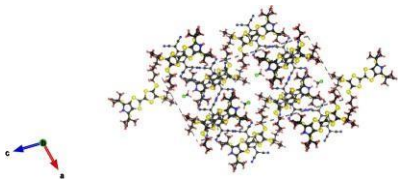


(a)Harvesteddriedpods(b)Sievedpowder

Figure2:Processingsequenceof*Albizia lebbeck*bio-adsorbent: (a) harvested and sun-dried pods prior to pulverization; and (b) groundand sieved bio-adsorbent powder (<425 µm) measured on an analytical balance.

Figure.3 presents the comprehensive flowchartillustratingthemethodologyadopted in this study, encompassing bio-adsorbent preparation, characterization, batch adsorption experiments,andAImodeldevelopment.Each step is discussed in detail in the subsequent sections. Thegeneralpropertiesandmolecular structure of methylene blue used in this study are summarized in Table 1.

Table 1: General characteristics and chemical structure of Methylene Blue (18).

| Characteristic | Value/Structure |
|---|---|
| Molecular formula | $C_{16}H_{18}ClN_3S$ |
| Molecular weight | 319.8 g $mol^{-1}$ |
| $\lambda_{max}$ | 667nm |
| Dye class | Thiazine(cationic) |
| Solubility in water | Soluble |
| **ChemicalStructure(2D)** | |
| |  |
| **3DMolecularRepresentation** | |
| |  |

## 3. Characterization

### PointofZeroCharge($pH_{PZC}$)

The surface charge characteristics of the*Albizia lebbeck* bio-adsorbent were evaluated by determining the point of zero charge ($pH_{PZC}$) using the pH drift method. The $pH_{PZC}$is a critical parameter governing the electrostatic interactions between the adsorbent surface and ionic pollutants, defined asthe pH at whichthe net surface charge of the adsorbent is zero. In this procedure, the initial pH of a 0.01 mol $L^{-1}$NaClelectrolytesolutionwasadjustedoverthe range2–12,andthefinalpHwasmeasuredafter reachingequilibrium.Aplotof($pH_i - pH_f$)ver-

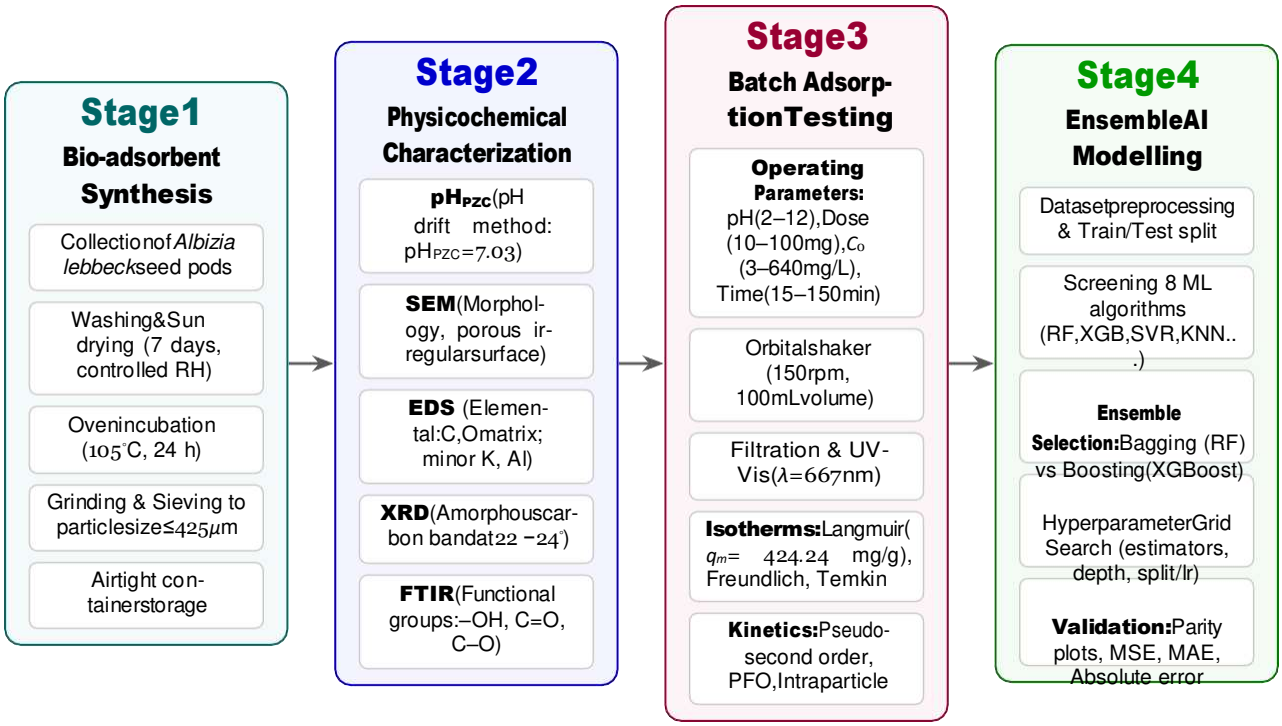



Figure 3: Comprehensive methodology flowchart illustrating the four sequential experimental and computational phases:(1) Bio-adsorbent synthesis and processing from*Albizia lebbeck* pods;(2)Physicochemicalcharacterization;(3)Batchadsorptionexperimentswithisothermand kineticevaluation;and(4)Ensemblemachinelearningmodellingandperformanceoptimization.

sus initial pH was constructed, and the $pH_{PZC}$ was identified as the point at which the curve intersected the zero line. The $pH_{PZC}$of the *Albizia lebbeck* bio-adsorbent was determined to be 7.03 (Figure 11).

## ScanningElectronMicroscopywith Energy Dispersive Spectroscopy (SEM–EDS)

The surface morphology and elemental composition of the synthesized bio-adsorbent were examined using SEM and EDS analysis. The SEM micrographs revealed a rough, non-uniform, and porous surface characterized by particlesofvaryingshapesandsizesdistributed across the surface, along with cavities and agglomerated structures. The abundant surface defects and extensive porous network indicate the availability of sufficient active sites conducivetodyemoleculeadsorption. Thisporous architectureisexpectedto enhancemass transfer and improve adsorbate accessibility to adsorption sites.

The EDS spectrum confirmed that the bio-adsorbent ispredominantlycomposedofcarbon (C) and oxygen (O), consistent with a carbonaceous matrix enriched with oxygen-containing functionalgroups.Minorquantitiesofpotassium (K) and aluminium (Al) were also detected, attributable to naturally occurring mineral constituents retained from the plant precursor. The high carboncontentprovidesthe structural framework of the adsorbent, while the oxygen-containing functional groups serveasactivesitesfacilitating adsorption through electrostatic interactions, hydrogen bonding, and surface complexation mechanisms (19).

## X-rayDiffraction(XRD)

X-ray diffraction analysis was employed to investigate the crystallographic structure of the bio-adsorbent.TheXRDpattern(Figure8)

exhibited a broad diffraction peak in the 2$\vartheta$range of 22–24°, indicative of a predominantly amorphous carbon framework with a significant degree of structural disorder.The diffuse, broad nature of this reflection confirms the absence of long-range crystallographic order. Furthermore, the lack of sharp diffraction peaks substantiates that the material possesses a predominantly amorphousstructurewithrandomorientations of aromatic layers and structural defects (20). This amorphous character is advantageous for adsorptionapplications,asitcreatesmoreac-tive sites and surface defects available for inter-action with adsorbate molecules.

*Itisworthnotingthattheamorphousna-tureofbio-adsorbentsisgenerallypreferredover highly crystalline structures for adsorption ap-plications, as disorder introduces a greater den-sity of edge sites, dangling bonds, and defect-associated functional groups that serve as high-energy adsorption sites.*

## FourierTransformInfraredSpec-troscopy (FTIR)

FTIR spectroscopy was performed over the wavenumber range of 4000–500 $cm^{-1}$to iden-tify the surface functional groups present on the bio-adsorbent. The principal absorption bands and their assignments are summarizedin Table 2.

These functional groups confirm the lig-nocellulosic nature of the bio-adsorbent and demonstrate the presence of abundant oxygen-containing active sites capable of participating in dye adsorption through multiple interaction mechanisms,includinghydrogenbonding,elec-trostatic attraction, and $\pi$–$\pi$interactions.

## BatchAdsorptionExperiments

Batch adsorption studies were systematically conductedtoevaluatetheremovalefficiencyof

Table 2: FTIR band assignments for*Albizia lebbeck* bio-adsorbent.

| **Peak($cm^{-1}$)** | **FunctionalGroupAssignment** |
|---|---|
| 3368.8 | O–Hstretching(cellulose,lignin, moisture) |
| 2930.9 | AsymmetricC–Hstretching(aliphatic) |
| 2126.6 | C≡CorC≡Nstretching |
| 1635.7 | C=Ostretching/H–O–Hbending |
| 1525.8 | AromaticC=Cskeletal(lignin) |
| 1374.3 | C–Nstretching |
| 1246.0 | C–Ostretching(phenolic) |
| 1100.4 | C–Ostretching(polysaccharides) |
| 510.9 | Mineralbendingvibrations |

the*Albizia lebbeck*pod-derived bio-adsorbent forMBfromaqueousmedia.Astocksolutionof MBwaspreparedanddilutedtoobtainworking concentrations in the range of 32–320 mg $L^{-1}$. The influence of key operational parameters—solutionpH(3–10),contacttime(15–150min), andadsorbentdosage (0.5g per100mL)—was investigatedunder controlledlaboratory condi-tions.

Experimentswere performed in250mLErlen-meyerflaskscontaining100mLofdyesolution, agitated at a constant speed of 150 rpm using an orbital shaker to ensure uniform adsorbate–adsorbent interaction. At predetermined time intervals,sampleswerecollectedandseparated by filtration using Whatman filter paper. The residual dye concentration was determined spec-trophotometrically at $\lambda_{max}$= 667 nm using a UV–visible spectrophotometer, with reference to a calibration curve established over the se-lected concentration range.

Figure4presents photographic documenta-tionofthebatchadsorptionexperimentalsetup, depicting dye solutions of varying concentra-tions and the gravity filtration assembly used for phase separation.

The adsorption performance was quantified intermsof percentageremoval ($\%R$) andequi-librium adsorption capacity ($q_e$) using the fol-



(a)MBsolutions (b)Gravityfiltration

Figure4:Photographicdocumentationofbatch adsorption experiments: (a) Erlenmeyer flasks containingMethyleneBluesolutionsdisplaying concentration-dependent color gradients during batch testing; and (b) gravity filtration assembly for solid–liquid separation of spent bio-adsorbent from treated effluent.

lowingstandardexpressions:

$$\%R = \frac{C_o - C_e}{C_o} \times 100 \quad (1)$$

$$q_e = \frac{(C_o - C_e) \cdot V}{m} \quad (2)$$

where$C_o$and$C_e$(mg $L^{-1}$) denote the initial and equilibrium dye concentrations, respectively;$V$(L)isthevolumeofthedyesolution; and $m$ (g) is the mass of the bio-adsorbent.

The experimental parameters and their ranges are summarized in Table 3.

To gain insight into the adsorption mechanism and surface behaviour, the equilibrium datawereanalysedusingLangmuir,Freundlich, andTemkinisothermmodels. Forkineticevaluation,adsorptiondatacollectedatdifferent

Table 3: Summary of experimental parameters investigated for MB adsorption onto the bio-adsorbent.

| No. | Parameter | Units | Low | High |
|---|---|---|---|---|
| 1 | pH | – | 2 | 12 |
| 2 | Adsorbentdose | mg | 10 | 100 |
| 3 | Initialconc. | $mgL^{-1}$ | 3 | 640 |

contact times were fitted to pseudo-first-order, pseudo-second-order, and intraparticle diffusion models.Model suitability was assessed using the coefficient of determination ($R^2$) and the nonlinear chi-square ($\chi^2$) test (21; 22).

## ArtificialIntelligenceModels

To accurately predict the adsorption performance ofthe *Albizia lebbeck* bio-adsorbent for MB removal, an ensemble machine learning framework was developed using Random Forest (RF) and eXtreme Gradient Boosting (XGBoost). Unlike conventional empirical regressiontechniquesthat oftenfailtocapturehighly nonlinear adsorption behaviour, ensemble learningalgorithms combinemultipledecisiontrees tomodelcomplexinteractionsamongexperimentalvariableswhilemaintainingexcellent predictive stability (23).

Thepredictiveframeworkutilizedexperimentallymeasuredadsorptionparameters—including solution pH, adsorbent dosage, contact time, and initial dye concentration—as input features, with percentage removal and equilibrium adsorption capacity ($q_e$) as target variables. Prior to model development, the experimental dataset was inspected for inconsistencies, redundant records, and missing observations to ensure data integrity.

*This investigation deliberately focused on two ensemble models representing fundamentallydifferent ensemble learning philosophies: Random Forestembodiesthebagging(bootstrapaggrega-*

*tion) paradigm, while XGBoost represents the boostingparadigmwithregularizedoptimization. This dual-model framework enables a comprehensive comparativeevaluationwhileavoiding unnecessary model redundancy.*

Sincetree-basedensemblealgorithmsareinherently insensitive to feature scaling, no normalization procedure was required prior to model training. The cleaned dataset was divided into training and testing subsets to evaluate model performance under unseen experimental conditions.

*RandomForest(RF)*

RandomForestisasupervisedensemblelearn-ing algorithm that generates predictions by aggregating the outputs of multiple decision trees constructed through bootstrap aggregation (bagging) (24). Each tree is trained on a randomly selected subset of both observations and predictor variables at each split, introducing diversity among individual trees, reducing theircorrelation,andimprovingtheoverallgeneralizationcapabilityoftheensemble.Thefinal prediction isobtained byaveraging theoutputs of all constituent trees, thereby substantially reducingpredictionvariance,minimizingoverfitting, and enhancing model robustness.

Figure5illustrates the architectural schematicoftheRandomForestmodelandthe decision tree ensemble mechanism implemented in this study.

*eXtremeGradientBoosting(XGBoost)*

XGBoostisanadvancedensemble learningalgorithm based on the gradient boosting principle, wherein decision trees are constructed sequentiallytominimizetheresidualerrorsofprecedingiterations(25).UnlikeRF,whichgenerates independent trees, each successive tree in XGBoost focuses on correcting the prediction errors of the previous tree, thereby progressively improving overall accuracy.

A key distinguishing feature ofXGBoost is its incorporation of L1 and L2 regularization penalties, which effectively control modelcomplexity, enhance generalization capability, and reduce overfitting risk (26). Additionally, XGBoost employs optimized tree construction, efficient pruning strategies, and second-order gradient optimization,enablingrapidconvergencewhile maintaining high predictive robustness.

Figure 6illustrates the sequential gradient boosting process and regularized additive learning framework of the XGBoost model.

*ModelEvaluationCriteria*

The predictive performance of both ensemble modelswasassessedusingseveralcomplementary statistical metrics:

- **Coefficientofdetermination**($R^2$):quantifiestheproportionofvarianceexplained by the model.
- **Mean squared error(MSE)** and **Root meansquarederror(RMSE)**:measure themagnitudeof predictionerrors.
- **Mean absolute error (MAE)**: provides the average absolute deviation between predicted and observed values.
- **Mean absolute percentage error (MAPE)**: offers a normalized, scale-independent measure of prediction accuracy.
- **Relative root mean squared error (RRMSE)**: enables comparison of predictionaccuracyrelativetotheobservedexperimental range.

Themathematicaldefinitionsofthekeyevaluation metrics are:

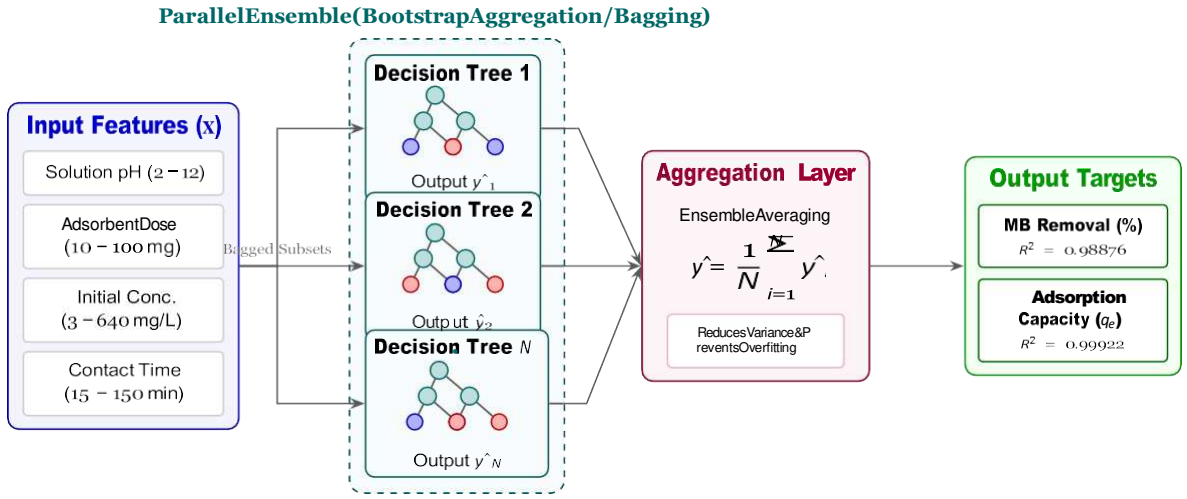



Figure 5: Architectural structure of the Random Forest model for predicting MB adsorption performance,illustratingfeatureinput,paralleldecisiontreegenerationviabootstrapaggregation (bagging), ensemble averaging, and dual target output prediction.

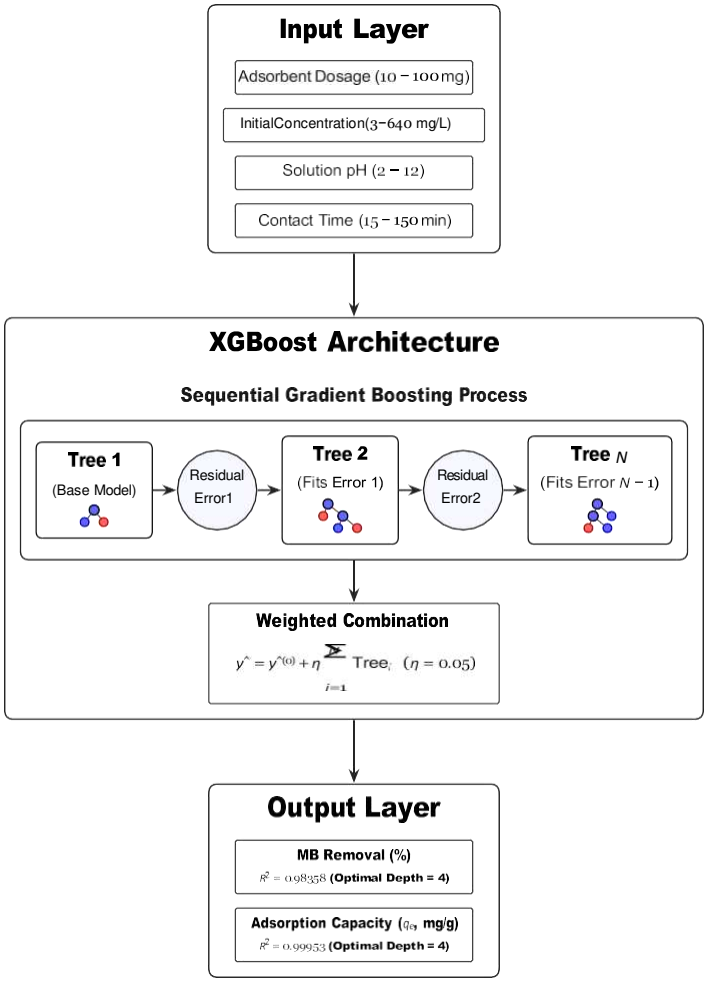



Figure6:ArchitecturalstructureoftheeXtremeGradientBoosting(XGBoost)model,detailing sequential tree construction, residual error minimization, regularized objective function, and final weighted response prediction.

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \quad (3)$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \quad (4)$$

$$\text{MAPE}=\frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i-\hat{y}_i}{y_i}\right| \quad (5)$$

where $y_i$and $\hat{y}_i$are theobserved and predicted values,respectively;$\bar{y}$ isthemeanofobserved values; and $n$ is the number of observations.

# 4. ResultsandDiscussion

## SEM–EDSAnalysis

TheSEMmicrographs(Figure7)revealedaheterogeneous surface morphology characterized by irregularly shaped particles, agglomerated structures, and an extensive network of pores distributed across the material surface. The rough, porous texture indicates the formation of a high-surface-area material, which is beneficial for adsorption processes as it provides abundantactivesitesfordyemoleculebinding.
The presence of cavities and surface defects creates favourable diffusion pathways for the transport of adsorbate molecules from the bulk solution to internal adsorption sites, thereby enhancing the overall adsorption performance.

The EDS spectrum of the selected region confirmed that carbon (C) and oxygen (O) are the predominant elements, consistent with the formation ofacarbonaceous matrix enriched with oxygen-containing functional groups. These oxygenated functionalities play a critical rolein the adsorption process by facilitating electrostatic interactions, hydrogen bonding, and surface complexation with dye molecules. Minor concentrations of potassium (K, 5.58 ppm) and aluminium (Al, 52.11 ppm) were detected, attributableto naturally occurring mineral constituents retained during sample preparation. The absence of significant impurity peaks confirms the synthesis of a high-purity adsorbent material (19).

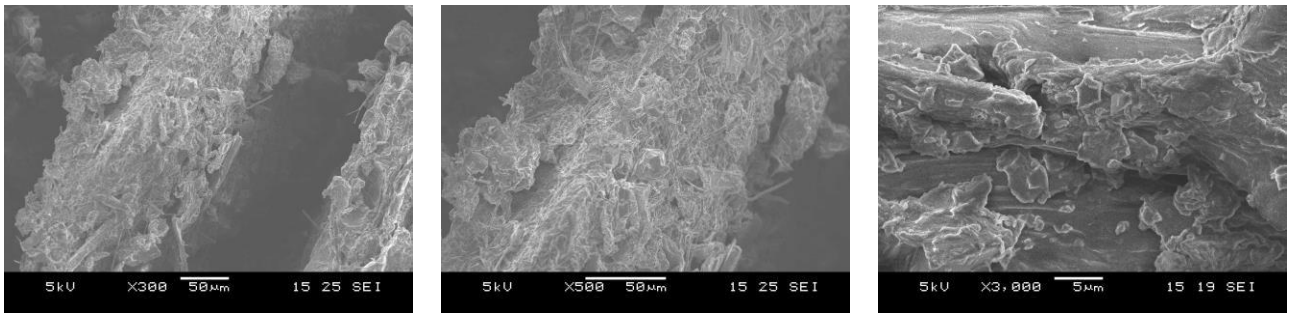
*ThehighC/OratioobservedintheEDS spectrumisconsistentwithapartiallycarbonizedlignocellulosicmatrix,wherethethermaltreatmentpreservedtheoxygen-containing functional groups essential for MB adsorption while establishing a robust carbon backbone.*

## XRDAnalysis

XRD analysis (Figure 8) confirmed that the synthesizedbio-adsorbentconsistspredominantlyof amorphous carbon with minimal structural ordering.Thebroaddiffractionpeakobserved inthe2$\vartheta$rangeof22–24°ischaracteristicof the irregular stacking of carbon layers and a highly disordered arrangement of carbon atoms. This structural disorder is advantageous for adsorption applications, as it generates additional active sites and surface defects capable of participatingintheadsorptionprocess.Theabsence of sharp crystalline peaks indicates minimal crystalline impurities and confirms the successful conversion of the lignocellulosic precursor into a carbon-rich adsorbent through the preparation process.

## FTIRAnalysis

The FTIR spectrum (Figure9) confirms the presence of diverse oxygen-containing functional groups on the bio-adsorbent surface that play significant roles in MB adsorption.The broad hydroxyl band centred at 3368.8 $cm^{-1}$indicates an abundance of polar functional groups that enhance adsorption affinity. The bands at 1635.7 $cm^{-1}$and 1525.8 $cm^{-1}$correspond to carbonyl and aromatic structures derived from lignin, which contribute to strong $\pi$–$\pi$ interactionswiththearomaticringsofMB molecules.Peaksinthe1374.3–1100.4$cm^{-1}$regiondemonstratethepresenceofcellulose-and hemicellulose-derived functional groups that increasethedensityofadsorptionsites.Overall,theFTIRanalysisconfirmsthatthenat-




(a)SEMat300×(50µm) (b)SEMat500×(50µm) (c)SEMat3000×(5µm)

Figure7:Scanningelectronmicrographs(SEM)ofthesynthesized*Albizialebbeck*bio-adsorbent at varying magnifications illustrating the porous, irregular, and heterogeneous surface texture conducive to dye adsorption.

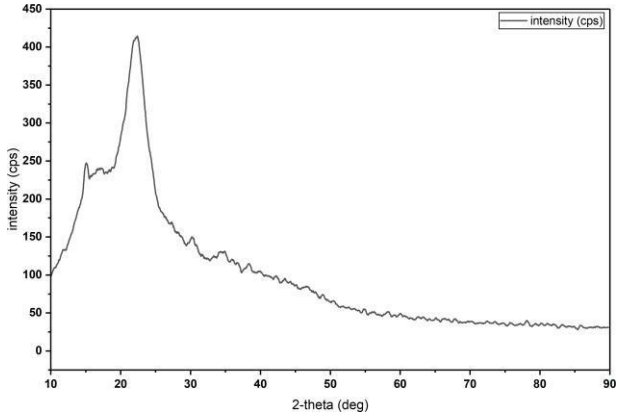



Figure 8: X-ray diffraction (XRD) pattern of the *Albizia lebbeck* bio-adsorbent showing the characteristicbroadpeakindicativeofanamorphous carbon framework.

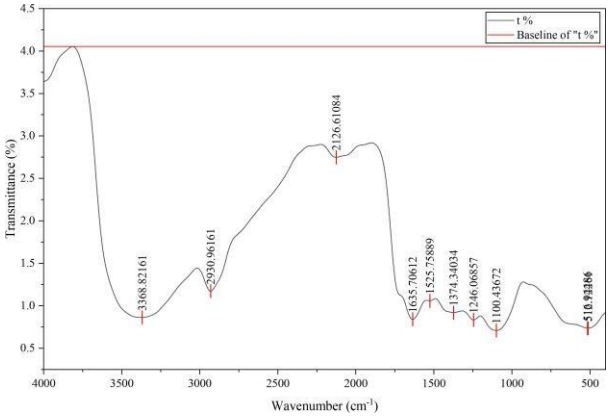



Figure 9: FTIR spectrum of the*Albizialebbeck*bio-adsorbent recorded between 4000 and500cm⁻¹identifyingkeysurfacefunctional groups.

urallyoccurringfunctionalgroupsonthebio-adsorbent are primarily responsible for the effective removal of MB from aqueous solution.

### AdsorptionIsothermAnalysis

Theequilibriumadsorptioncharacteristicswere investigatedbyfittingtheexperimentaldata toLangmuir,Freundlich,andTemkinisotherm models.The isotherm parameters and statistical indicatorsarepresentedinTable4,andthe corresponding nonlinear model fits are shown in Figure 10.

The nonlinear forms of the isotherm models employed are:

**Langmuir:**

$$q_e = \frac{q_m K_L C_e}{1 + K C_{Le}} \quad (6)$$

**Freundlich:**

$$q_e = K_F \quad C_e^{1/n} \quad (7)$$

**Temkin:**

$$q_e = B \ln(K_T C_e) \quad (8)$$

where$q_m$(mgg⁻¹)isthemaximummonolayer adsorptioncapacity; $K_L$(Lmg⁻¹) is the Langmuir constant; $K_F$and $n$are Freundlich con-

stants;$K_T$(mgL$^{-1}$)istheTemkinequilibrium binding constant; and $B = RT/b_T$(Jmol$^{-1}$) is related to the heat of adsorption.

The Langmuir model yielded a good correlationcoefficient($R^2$=0.9635),suggestingsignificant monolayer coverage as active sites are progressivelyoccupied(27).Thedimensionlessseparationfactor$R_L$= 0.6896fallswithin the range 0< $R_L$<1, confirming that the adsorptionprocessisfavourableundertheexperimentalconditionsinvestigated.Themaximummonolayeradsorptioncapacitywasdetermined to be 424.24 mg g$^{-1}$, which is notably competitive compared with many reported bioadsorbents.

The Freundlich model exhibited a comparatively lower$R^2$value of 0.8186 (linear form), indicating that it is less suited to describe the equilibriumdata.Nevertheless,the Freundlich exponent $n$ = 1.0207 (where $n$ >1) confirms favourable adsorption.

The Temkin model provided the best overall fit ($R^2$= 0.9733), indicating that adsorbate–adsorbentinteractionsandadsorptionenergy varywithincreasingsurfacecoverage(28).The strongTemkincorrelation,coupledwiththeexcellent Langmuir fit, suggests that both monolayer adsorption and energy-dependent surface interactions contribute significantly to the uptake of MB by the bio-adsorbent.

## pH$_{PZC}$Analysis

The determined pH$_{PZC}$of 7.03 indicates that the bio-adsorbent surface undergoes a charge transition from positive to negative at approximately neutral pH. At solution pH values below 7.03, the adsorbent surface carries a net positivecharge,resultinginelectrostaticrepulsion with the cationic MB molecules and consequently reduced adsorption efficiency.Conversely, at pH values above 7.03, progressive deprotonation of hydroxyl, carboxyl, and phenolic functional groups generates a negatively chargedsurfacethatstronglyattractspositively chargedMBmoleculesthroughelectrostaticinteractions, thereby enhancing adsorption.

## AdsorptionKinetics

The adsorption kinetics were evaluated using pseudo-first-order (PFO), pseudo-second-order (PSO), and Weber–Morris intraparticle diffusion models. The kinetic parameters are presentedinTable5,andthecorrespondingmodel fits are shown in Figure 12.

The linearized kinetic model equations employed are:

**Pseudo-first-order:**

$$\ln(q_e - q_t) = \ln q_e - k_1 t \quad (9)$$

**Pseudo-second-order:**

$$\frac{t}{q_t} = \frac{1}{k_2 q_e^2} + \frac{t}{q_e} \quad (10)$$

**Intraparticlediffusion:**

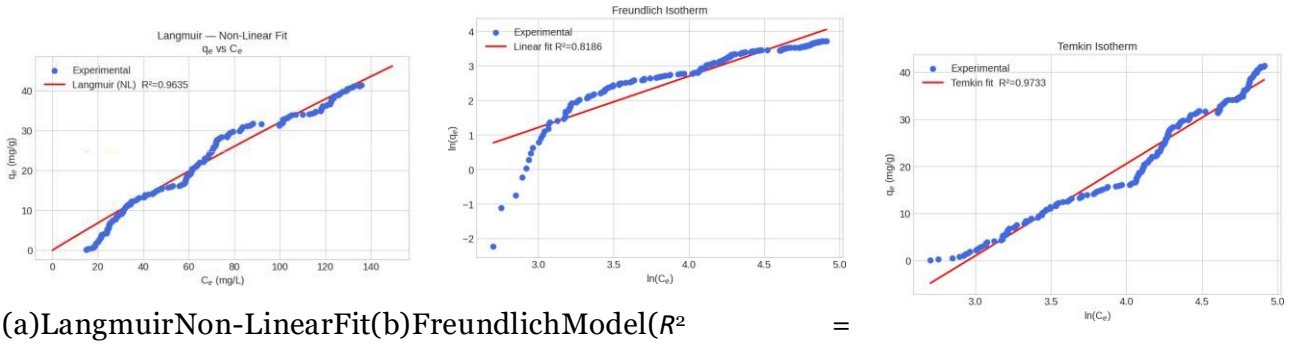
$$q_t = k_d t^{0.5} + C \quad (11)$$

where$q_t$(mg g$^{-1}$)istheadsorptioncapacity attime$t$;$k_1$(min$^{-1}$),$k_2$(g mg$^{-1}$min$^{-1}$),and $k_d$(mg/g/min$^{0.5}$) are the rate constants for the respective models; and $C$(mgg$^{-1}$) is the intercept related to boundary layer thickness.

ThePFOmodelexhibitedamoderatefit ($R^2$= 0.9119), indicating its inability to satisfactorilydescribetheoveralladsorptionprocess. In contrast, the PSO model demonstrated a significantlyhigher$R^2$valueof0.9632,confirming excellent agreement with the experimental data and indicating that the adsorption process

Table4:Isothermmodelparameters forMBadsorptiononto*Albizialebbeck*bio-adsorbent.

| Model | Parameter | Value | Unit |
|---|---|---|---|
| Langmuir | $q_m$ | 424.2415 | mgg$^{-1}$ |
| | $K_L$ | 0.000818 | Lmg$^{-1}$ |
| | $R_L$ | 0.6896 | – |
| | $R^2$ | 0.9635 | – |
| | $\chi^2$ | 82.3113 | – |
| Freundlich | $K_F$ | 0.3521 | – |
| | $n$ | 1.0207 | – |
| | $R^2$ | 0.9614 | – |
| | $\chi^2$ | 81.0917 | – |
| Temkin | $K_T$ | 0.0524 | mgL$^{-1}$ |
| | $B$ | 19.5660 | Jmol$^{-1}$ |
| | $R^2$ | 0.9733 | – |
| | $\chi^2$ | 115.2589 | – |




(a)LangmuirNon-LinearFit($R^2$=0.9635) (b)FreundlichModel($R^2$=0.8186) (c)TemkinModel($R^2$=0.9733)

Figure10:AdsorptionisothermfitsforMBuptakeonto*Albizialebbeck*bio-adsorbent:(a) Langmuir non-linear model, (b) Freundlich isotherm, and (c) Temkin isotherm.

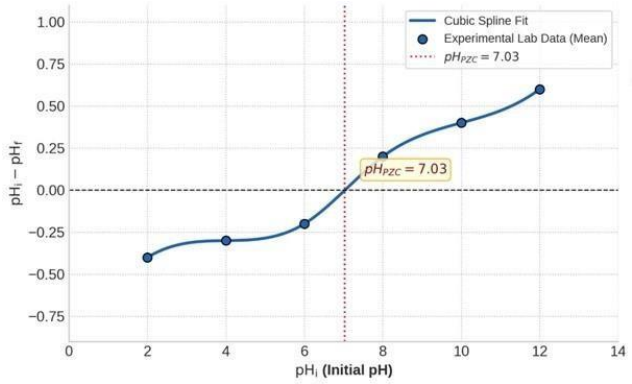



Figure 11: Determination of the point of zero charge ($pH_{PZC}$= 7.03)ofthe *Albizia lebbeck* bio-adsorbentby thepHdriftmethodwithcubic splinefit.

Table5:KineticmodelparametersforMB adsorptiononto*Albizialebbeck*bio-adsorbent.

| Model | Param. | ValueUnit |
|---|---|---|
| Pseudo-first-order | $R^2$ | 0.9119– |
| | $k_1$ | −0.0167min$^{-1}$ |
| | $q_e$ | 2.41mgg$^{-1}$ |
| Pseudo-second-order | $R^2$ | 0.9632– |
| | $k_2$ | −0.002214gmg$^{-1}$min$^{-1}$ |
| | $q_e$ | 23.31mgg$^{-1}$ |
| Intraparticle diffusion | $R^2$ | 0.9785– |
| | $k_d$ | −2.8572mg/g/min$^{0.5}$ |
| | $C$ | 60.47mgg$^{-1}$ |

ispredominantlygovernedbychemisorption—involving interactions between MB molecules andtheactivefunctionalgroupsontheadsor-bent surface.

The Weber–Morris intraparticle diffusion modelyieldedhighlinearity($R^2$=0.9785);

however,thediffusionplotdidnotpassthrough the origin and exhibited a considerable intercept ($C$ = 60.47 mg $g^{-1}$), indicating that intraparticlediffusionisnotthesolerate-controlling step. Instead, the adsorption process involves multiple mass-transfer mechanisms, including external film diffusion followed by intraparticle diffusion within the porous structure of the adsorbent.

## MachineLearningModelDevelopment and Performance

*ComparativeModelScreening*

Eight machine learning algorithms were initially screened to identify the most suitable models for predicting MB removal efficiency and adsorption capacity. The comparative$R^2$scores are presented in Table 6.

Table 6: Comparative performance ($R^2$)of machine learning models for MB adsorption prediction.

| Model | MBRemoval | $q_e$ |
|---|---|---|
| RandomForest | 0.9853 | 0.9991 |
| XGBoost | 0.9560 | 0.9994 |
| GradientBoosting | 0.9620 | 0.9992 |
| LightGBM | 0.9510 | 0.9989 |
| DecisionTree | 0.9584 | 0.9968 |
| ExtraTrees | 0.9240 | 0.9910 |
| SVR | 0.0137 | 0.9950 |
| KNN | −2.899 | 0.8610 |

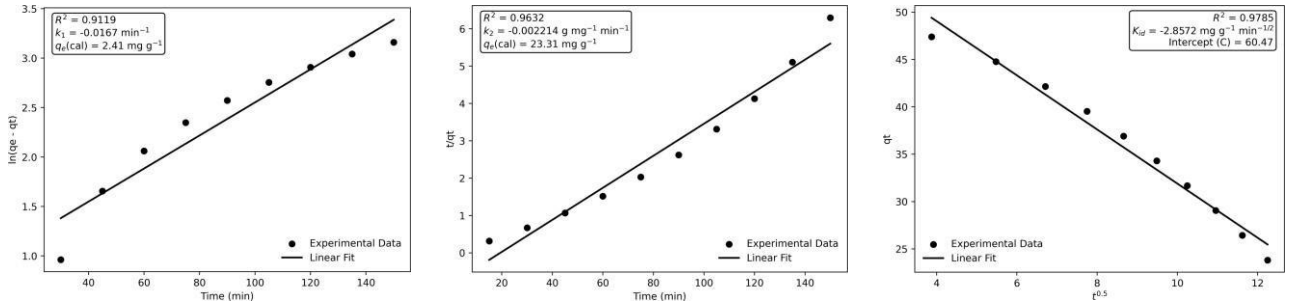
Amongallinvestigatedmodels,RandomFor-est demonstrated the highest predictive accu-racy for percentage removal ($R^2$= 0.9853), while both RF and XGBoost achieved excep-tional prediction performance for adsorption capacity ($R^2$>0.999).Notably, SVR andKNN exhibited poor predictive performance forpercentageremoval,withKNNproducinga negative$R^2$value—indicatingpredictionsworse than a simple mean baseline. This poor performance can be attributed to the inherent limitations of distance-based algorithms (KNN) and kernel-based methods (SVR) when applied to relatively small, high-dimensional adsorption datasets with complex nonlinear interactions. Consequently,RFwasselectedastheproposed predictive model, and XGBoost was employed as the comparative benchmark.

*HyperparameterOptimization*

Systematic hyperparameter optimization was performed for both RF and XGBoost models. Thecompleteoptimizationresultsarepresented in Tables 7–10.

**RandomForestOptimization.** FortheRF model, increasing the maximum tree depth from3to7progressivelyenhancedpredic-tive performance for both target variables, reflected by higher $R^2$values and corresponding reductions in MSE, RMSE, MAE, MAPE, and RRMSE.Theoptimalconfiguration—200estimators, maximum depth of 7, and minimum samplessplitof2—yieldedan$R^2$of **0.98876** for MB removal and **0.99922** for adsorption capacity.Furtherincreasingthedepthto8 did not improve performance and introduced slightly higher prediction errors, indicating the onset of overfitting.

**XGBoost Optimization.**For XGBoost, the optimal hyperparameter combination—200 estimators, maximum depth of 4, and learning rate of 0.05—achieved $R^2$values of **0.98358** for MB removal and **0.99953** for adsorption capacity. Notably, XGBoost reached optimal performanceatashallowertreedepth(4versus 7 for RF), which is consistent with the boostingparadigm wheresequential errorcorrection compensates for limited individual tree complexity.




(a) Pseudo-first-order($R^2$= 0.9119) (b) Pseudo-second-order($R^2$= 0.9632) (c) Intraparticlediffusion($R^2$= 0.9785)

Figure12:AdsorptionkineticmodelsforMBremovalonto*Albizialebbeck*bio-adsorbent:(a) Pseudo-first-order, (b) Pseudo-second-order, and (c) Intraparticle diffusion plots.

Table7:HyperparameteroptimizationresultsfortheRandomForestmodelpredicting%MB removal.

| $n_{est}$ | Depth | MinSpl | $R^2$ | MSE | RMSE | MAE | MAPE(%) |
|---|---|---|---|---|---|---|---|
| 100 | 3 | 2 | 0.94661 | 1.66044 | 1.28858 | 1.06308 | 1.44134 |
| 100 | 4 | 2 | 0.96586 | 1.06194 | 1.03051 | 0.81986 | 1.11319 |
| 100 | 5 | 5 | 0.98370 | 0.50688 | 0.71196 | 0.57534 | 0.79565 |
| 100 | 6 | 2 | 0.98706 | 0.40253 | 0.63446 | 0.53324 | 0.74096 |
| 100 | 7 | 5 | 0.98692 | 0.40690 | 0.63789 | 0.48942 | 0.68529 |
| 100 | 8 | 2 | 0.98529 | 0.45766 | 0.67651 | 0.52780 | 0.73340 |
| 200 | 3 | 5 | 0.94524 | 1.70305 | 1.30501 | 1.08243 | 1.47223 |
| 200 | 4 | 5 | 0.96527 | 1.08026 | 1.03935 | 0.83268 | 1.12972 |
| 200 | 5 | 2 | 0.98413 | 0.49361 | 0.70257 | 0.58731 | 0.80504 |
| 200 | 6 | 2 | 0.98748 | 0.38945 | 0.62406 | 0.53699 | 0.74221 |
| 200 | 7 | 2 | **0.98876** | **0.34968** | **0.59134** | **0.47317** | **0.65648** |
| 200 | 8 | 5 | 0.98609 | 0.43249 | 0.65764 | 0.50411 | 0.70201 |

*RandomForestPredictionPerformance*

Theparityplots(Figure13)comparetheexperimentally measured and RF-predicted values across all evaluated hyperparameter configurations. The prediction points are closely distributed around the 45°parity line, indicating excellent agreement between experimental observationsand modelpredictions.The minimal deviationdemonstratesthattheRFmodeleffectively captured the nonlinear relationships betweenadsorptionvariablesandresponseparameters while maintaining stable prediction performanceacrosstheentireexperimental range.

*XGBoostPredictionPerformance*

The XGBoost parity plots (Figure 14) demonstrate closely aligned prediction points along the 45°parity line for both target variables, confirming that the gradient boosting framework effectively learned the nonlinear adsorption behaviour while maintaining robust predictive performance. The optimized XGBoost configurationachieved$R^2$valuesof0.98358and 0.99953forMBremovalandadsorptioncapac-ity, respectively.

*ErrorAnalysisoftheDevelopedModels*

Figure15presents the variation in mean squarederror(MSE)oftheoptimizedRan-

Table8:HyperparameteroptimizationresultsfortheRandomForestmodelpredictingadsorption capacity ($q_e$).

| $n_{est}$ | Depth | MinSpl | $R^2$ | MSE | RMSE | MAE | MAPE(%) |
|---|---|---|---|---|---|---|---|
| 100 | 3 | 2 | 0.99371 | 0.90805 | 0.95292 | 0.77365 | 7.93691 |
| 100 | 4 | 2 | 0.99870 | 0.18711 | 0.43256 | 0.34334 | 3.19272 |
| 100 | 5 | 5 | 0.99920 | 0.11601 | 0.34060 | 0.26997 | 2.19323 |
| 100 | 6 | 2 | 0.99932 | 0.09885 | 0.31440 | 0.25171 | 2.04581 |
| 100 | 7 | 5 | 0.99923 | 0.11110 | 0.33331 | 0.25442 | 1.92815 |
| 100 | 8 | 2 | 0.99908 | 0.13220 | 0.36360 | 0.27140 | 2.05269 |
| 200 | 3 | 5 | 0.99366 | 0.91584 | 0.95699 | 0.77843 | 8.06282 |
| 200 | 4 | 5 | 0.99877 | 0.17713 | 0.42087 | 0.34237 | 3.28395 |
| 200 | 5 | 2 | 0.99918 | 0.11851 | 0.34425 | 0.27324 | 2.24344 |
| 200 | 6 | 2 | 0.99927 | 0.10533 | 0.32454 | 0.25647 | 2.07754 |
| 200 | 7 | 2 | **0.99922** | **0.11283** | **0.33591** | **0.25490** | **2.01771** |
| 200 | 8 | 5 | 0.99913 | 0.12558 | 0.35437 | 0.25991 | 2.02268 |

Table9:HyperparameteroptimizationresultsfortheXGBoostmodelpredicting%MBremoval.

| $n_{est}$ | Depth | Learn.Rat | $R^2$ | MSE | RMSE | MAE | MAPE(%) |
|---|---|---|---|---|---|---|---|
| 100 | 3 | 0.05 | 0.97741 | 0.70250 | 0.83816 | 0.63623 | 0.90424 |
| 100 | 4 | 0.05 | 0.98622 | 0.42868 | 0.65474 | 0.50846 | 0.71502 |
| 100 | 5 | 0.01 | 0.77431 | 7.01957 | 2.64945 | 1.91000 | 2.75632 |
| 100 | 6 | 0.05 | 0.98028 | 0.61322 | 0.78309 | 0.56105 | 0.79678 |
| 100 | 7 | 0.05 | 0.97414 | 0.80438 | 0.89687 | 0.64480 | 0.90613 |
| 100 | 8 | 0.10 | 0.96214 | 1.17760 | 1.08517 | 0.81149 | 1.12265 |
| 200 | 3 | 0.01 | 0.94708 | 1.64593 | 1.28294 | 1.00743 | 1.35552 |
| 200 | 4 | 0.05 | **0.98358** | **0.51059** | **0.71456** | **0.52857** | **0.74008** |
| 200 | 5 | 0.05 | 0.97882 | 0.65886 | 0.81170 | 0.61125 | 0.85133 |
| 200 | 6 | 0.05 | 0.97479 | 0.78425 | 0.88558 | 0.63079 | 0.88586 |
| 200 | 7 | 0.10 | 0.96979 | 0.93956 | 0.96931 | 0.68078 | 0.95383 |
| 200 | 8 | 0.10 | 0.95657 | 1.35070 | 1.16220 | 0.89675 | 1.23476 |

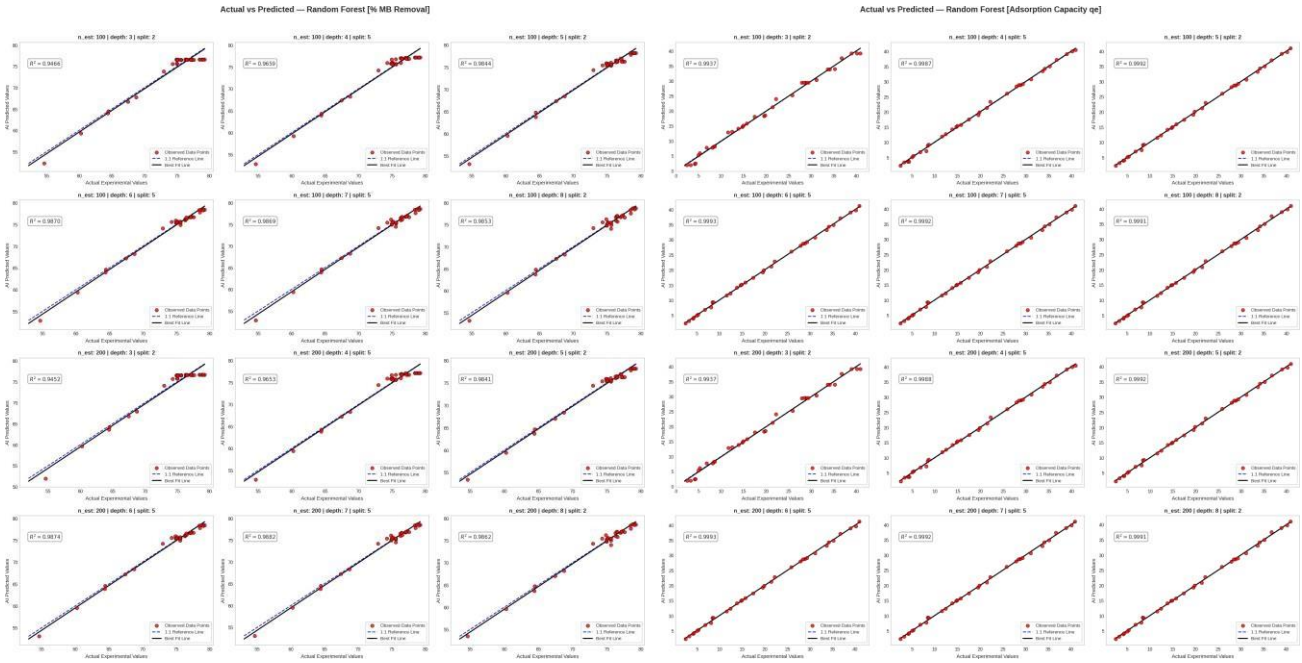
domForestandXGBoostmodelswithrespect tomaximumtreedepthwhilekeepingthere-maininghyperparametersfixed.Bothmodels exhibitedimproved prediction accuracywith increasing tree depth until the optimum con-figuration was reached.The Random Forest modelachievedtheminimumMSEatamax-imum depth of 7, whereas XGBoost attained its lowest prediction error at a maximum depthof4.Beyondtheseoptimaldepths,thepredi c-tionerrorincreased,indicatingnofurtherim-provementinmodelperformance.Theresults demonstratethatmaximumtreedepthplaysa significantroleincontrollingmodelcomplexity and prediction accuracy, thereby supporting theselected hyperparameter configurations for the final ensemble models.

*RelativeErrorAnalysis*

Figure16presents the sample-wise absolute error distributions of the Random Forest (RF) and XGBoost models for predicting % Methy-leneBlueremovalandadsorptioncapacity($q_e$). Absoluteerrorwasusedtodirectlyevaluatepre-diction accuracy and to avoid the limitations associated with percentage-based error mea-sures at lower response values.For % Methy-

Table 10: Hyperparameter optimization results for the XGBoost model predicting adsorption capacity ($q_e$).

| $n_{est}$ | Depth | Learn.Rat | $R^2$ | MSE | RMSE | MAE | MAPE(%) |
|---|---|---|---|---|---|---|---|
| 100 | 3 | 0.05 | 0.99942 | 0.08336 | 0.28873 | 0.23693 | 1.94756 |
| 100 | 4 | 0.05 | 0.99944 | 0.08028 | 0.28334 | 0.22023 | 1.71709 |
| 100 | 5 | 0.01 | 0.85665 | 20.70699 | 4.55049 | 4.00003 | 47.36721 |
| 100 | 6 | 0.05 | 0.99915 | 0.12225 | 0.34964 | 0.29037 | 2.62644 |
| 100 | 7 | 0.05 | 0.99893 | 0.15524 | 0.39400 | 0.33341 | 2.93611 |
| 100 | 8 | 0.10 | 0.99911 | 0.12792 | 0.35765 | 0.29783 | 2.60121 |
| 200 | 3 | 0.01 | 0.97860 | 3.09175 | 1.75834 | 1.54263 | 17.75000 |
| 200 | 4 | 0.05 | **0.99953** | **0.06778** | **0.26035** | **0.19772** | **1.66151** |
| 200 | 5 | 0.05 | 0.99945 | 0.07960 | 0.28213 | 0.23258 | 1.71099 |
| 200 | 6 | 0.05 | 0.99930 | 0.10079 | 0.31748 | 0.25422 | 2.14091 |
| 200 | 7 | 0.10 | 0.99905 | 0.13731 | 0.37055 | 0.30726 | 2.55694 |
| 200 | 8 | 0.10 | 0.99908 | 0.13269 | 0.36427 | 0.30161 | 2.59575 |

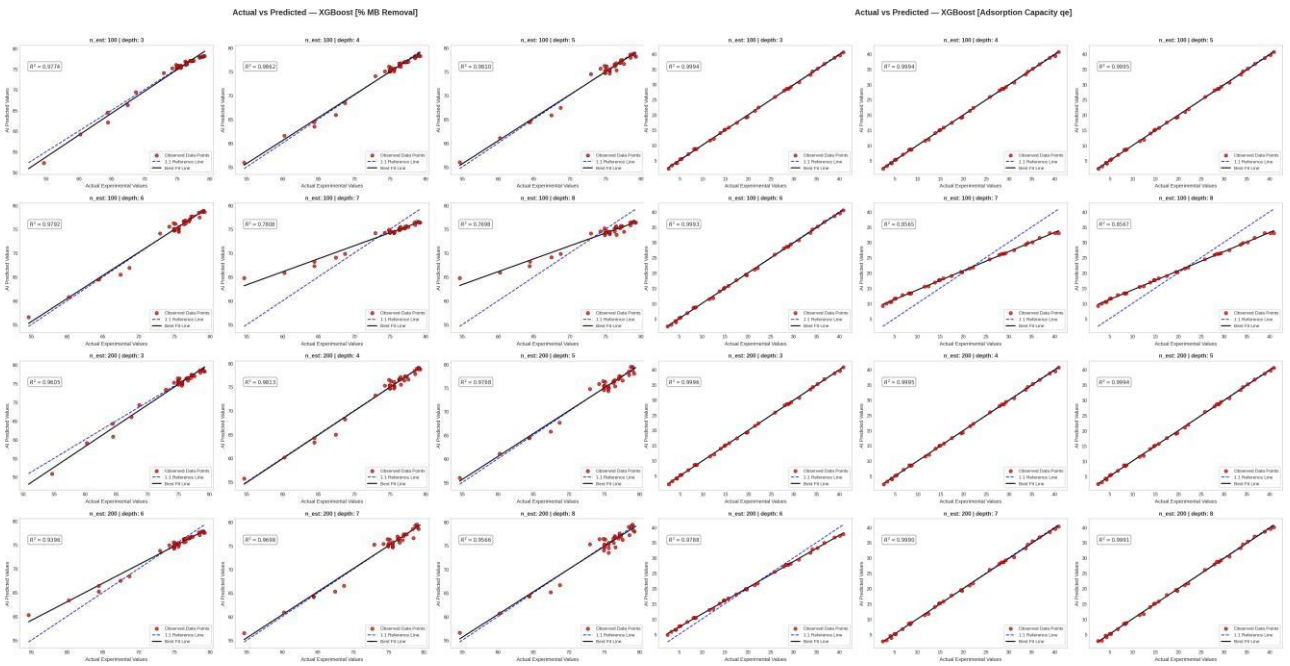



(a) Actualvs.Predicted:%MBRemoval (b)Actualvs.Predicted:AdsorptionCapacity($q_e$)

Figure 13: Parity plots comparing experimental observations with Random Forest model predictions across the evaluated hyperparameter combinations: (a) percentage removal of MB and (b) equilibrium adsorption capacity ($q_e$).

lene Blue removal, the Random Forest model exhibited a maximum absolute error of 1.39%, whereas XGBoost reached 10.14%.Similarly, for adsorption capacity ($q_e$), Random Forest maintained excellent prediction stability withamaximum absolute error of only 0.91mgg$^{-1}$, whileXGBoostshowedahighermaximumdeviationof7.84mgg$^{-1}$.Overall,theconsistently lower and more uniform error distribution of the Random Forest model indicates its superiorabilitytocapturethenonlinearadsorption behaviour compared with XGBoost.

The overall results demonstrate that ensemblelearningmodelsprovideaneffectiveandreliableframeworkforpredictingmethyleneblue adsorption using experimentally measured op-




(a) Actualvs.Predicted:%MBRemoval (b)Actualvs.Predicted:AdsorptionCapacity($q_e$)

Figure 14: Parity plots comparing experimental observations with XGBoost model predictions across the evaluated hyperparameter configurations:(a) percentage removal of MB and (b) equilibrium adsorption capacity ($q_e$).

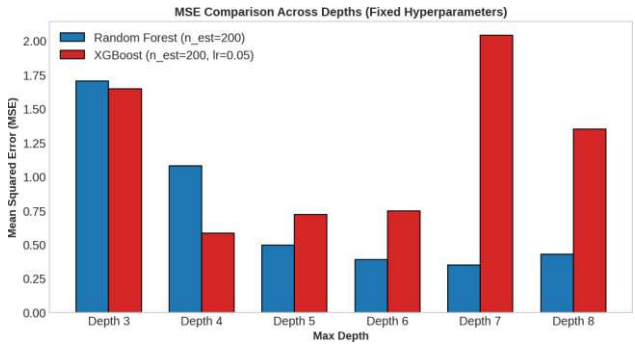



Figure15:MeanSquaredError(MSE)comparison between Random Forest ($n_{est}$= 200) and XGBoost ($n_{est}$= 200,lr = 0.05) across varying maximum tree depths (3 to 8).

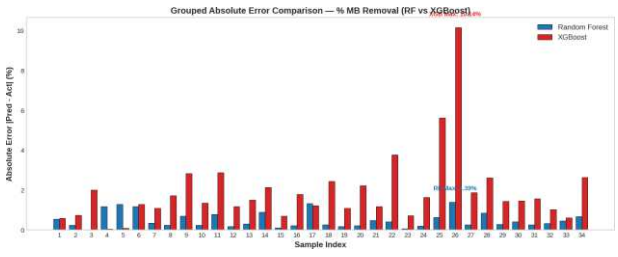
eratingparameters. BothRandomForestand XGBoost achieved excellent agreement with experimental observations. The optimized Random Forest model produced the highest prediction accuracy for percentage MB removal ($R^2$=0.98876),whereasXGBoostachievedа marginallyhigherpredictionaccuracyforadsorption capacity ($R^2$= 0.99953) compared with Random Forest ($R^2$= 0.99922). These resultsindicatethatbothensemblemodelsare highlyeffective,withRandomForestproviding thebest overall balancebetween predictionaccuracy, robustness, and generalization across both target variables.

## 5. Conclusion

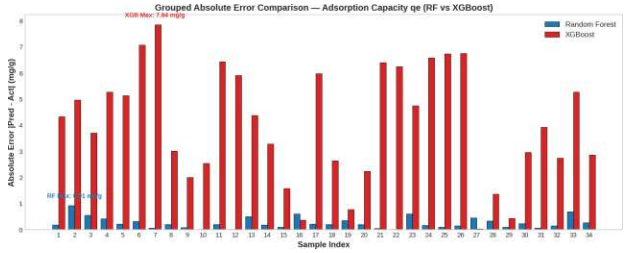
The present study demonstrated that*Albizia lebbeck*pod-derivedbio-adsorbent isanefficient, low-cost, and environmentally sustainable material for the removal of methylene blue from aqueous solution. The principal findings are summarized as follows:

(i) Comprehensive characterization by $pH_{PZC}$,SEM–EDS,XRD,andFTIR confirmed the formation of a porous, amorphouscarbonstructurewithahighconcentration ofoxygen-containing functional groups, providing abundant active sites for dye adsorption.

(ii) The adsorption equilibrium data were bestdescribedbytheLangmuirisotherm




(a) Sample-wiseAbsoluteError:%MBRemoval




(b) Sample-wiseAbsoluteError:AdsorptionCapacity ($q_e$)

Figure 16: Grouped absolute error distributions across individual experimental samples comparing Random Forest and XGBoost predictions for:(a) percentage removal of MB and (b) equilibrium adsorption capacity ($q_e$).

model,indicatingmonolayeradsorptionwithamaximumcapacityof 424.24mgg$^{-1}$.The Temkin model also provided an excellent fit ($R^2$= 0.9733), suggesting energy-dependent surface interactions.

(iii) Kinetic analysis confirmed that the pseudo-second-ordermodelbestrepresents the adsorption process ($R^2$= 0.9632), indicating that chemisorption throughsurfacefunctional groupinteractions constitutes the primary rate-limiting mechanism.

(iv) Among eight machine learning algorithmsevaluated, ensemble models demon-strated superior predictive performance. The optimized Random Forest model achieved $R^2$= 0.98876 for MB removal prediction,whiletheoptimizedXGBoost model attained $R^2$= 0.99953 for adsorption capacity prediction.

(v) Thecomplementarystrengthsofbagging-based (RF) and boosting-based (XGBoost) ensemble learning approaches were validated, demonstrating that machine learning can effectively model complex, nonlinear adsorption systems with high predictive accuracy.

*Future investigations should explore the application of deep learning architectures (e.g., neural networks with attention mechanisms)for adsorption prediction, the incorporation of additional input features such as temperature andionicstrength,andthedeploymentofthese modelsinreal-timewastewatertreatmentmonitoringsystems. Theintegrationofbio-adsorbent development with AI-driven process optimization holds significant promise for advancing scalable, sustainable wastewater remediation technologies.*

## References

[1] Schwarzenbach,R.P.,Escher,B.I.,Fenner, K., Hofstetter, T.B., Johnson, C.A., von Gunten, U., Wehrli, B., “The challenge of micropollutants in aquatic systems,”*Science*, vol. 313, no. 5790,pp. 1072–1077, 2006.

[2] Holkar,C.R.,Jadhav,A.J.,Pinjari,D.V., Mahamuni, N.M., Pandit, A.B., “A critical reviewontextile wastewater treatments: Recentadvancements,”*Journalof EnvironmentalManagement*,vol.182,pp.351–366,2016.

[3] Sultana, R., Mudliar, S., et al., “Environmentalimpactoftextileeffluentdischarge

and remediation technologies," *Water Research*, vol. 170, p. 115340, 2020.

[4] Hameed, B.H., Din, A.T.M., Ahmad, A.L., "Adsorption of methylene blue onto bamboo-based activated carbon: Kinetics and equilibrium studies," *Journal of HazardousMaterials*,vol.141,no.3,pp.819–825,2007.

[5] Rafatullah, M., Sulaiman, O., Hashim, R., Ahmad,A.,"Adsorptionofmethyleneblue on low-cost adsorbents: A review,"*Journalof HazardousMaterials*,vol.177,no.1–3, pp. 70–80, 2010.

[6] Oz, M., Lorke, D.E., Petroianu, G.A.,"MethyleneblueandAlzheimer'sdisease," *BiochemicalPharmacology*,vol.78,no.8, pp.927–932,2009.

[7] Tan, K.A., Morad, N., Teng, T.T., Norli,I.,Panneerselvam,P.,"Removalofcationic dyebyadsorptionontoprocessedoilpalm fronds,"*Desalination and Water Treatment*, vol. 51, pp. 2487–2499, 2013.

[8] Gupta, V.K., Suhas, "Application of low-costadsorbentsfordyeremoval:Areview," *Journal of Environmental Management*, vol. 90, no. 8, pp. 2313–2342, 2009.

[9] Crini, G., "Non-conventional low-cost adsorbentsfordyeremoval:Areview," *BioresourceTechnology*, vol.97,no.9, pp.1061–1085,2006.

[10] Ali,I.,"Newgenerationadsorbentsforwatertreatment,"*ChemicalReviews*,vol.112, no. 10, pp. 5073–5091, 2012.

[11] Chau,K.W.,"Areviewonintegrationofartificialintelligenceintowaterqualitymodelling,"*Marine Pollution Bulletin*, vol. 52, no. 6, pp. 726–733, 2006.

[12] Najafzadeh, M., Zeinolabedini, M., "Prognosticationofwaterqualityindexbyus-ing artificial intelligence methods," *Water Resources Management*, vol. 33, pp. 3959–3977,2019.

[13] LeCun,Y., Bengio,Y.,Hinton,G., "Deep learning,"*Nature*,vol.521,no.7553, pp.436–444,2015.

[14] Garg, A., et al., "Applications of artificial intelligenceandmachinelearninginwater and wastewater treatment: A comprehensive review," *Journal of Cleaner Production*, vol. 330, p. 129800, 2022.

[15] Melesse,A.M.,etal.,"Artificialneuralnetworkapplicationsinwaterresources,"*Sensors*, vol. 11, no. 4, pp. 4005–4018, 2011.

[16] Palani, S., Liong, S.Y., Tkalich, P., "An ANN application for water quality forecasting,"*Marine Pollution Bulletin*, vol. 56, no. 9, pp. 1586–1597, 2008.

[17] Bansal, M., Garg, U., Singh, D., Garg, V.K., "Removal of Cr(VI) from aqueous solutionsusing pre-treated*Albizia lebbeck* podbiomass,"*JournalofHazardousMaterials*,vol.167,no.1–3,pp.619–628,2009.

[18] O'Neil, M.J., *The Merck Index: An Encyclopedia of Chemicals, Drugs, and Biologicals*, 15th ed., Royal Society of Chemistry, Cambridge, UK, 2013.

[19] Foo, K.Y., Hameed, B.H., "Insights into the modeling of adsorption isotherm systems,"*Chemical Engineering Journal*, vol. 156, no. 1, pp. 2–10, 2010.

[20] Franklin,R.E.,"Crystallitegrowthin graphitizing and non-graphitizing carbons," *ProceedingsoftheRoyalSocietyofLondon A*, vol. 209, no. 1097, pp. 196–218, 1951.

[21] Ho, Y.S., McKay, G., “Pseudo-second order-model for sorption processes,” *Process Biochemistry*, vol. 34, no. 5, pp. 451–465, 1999.

[22] Weber, W.J., Morris, J.C., “Kinetics of adsorption on carbon from solution,” *Journal of the Sanitary Engineering Division*, vol. 89, no. 2, pp. 31–60, 1963.

[23] Dietterich, T.G., “Ensemble methods in machine learning,” in *Multiple Classifier Systems*, Springer, Berlin, pp. 1–15, 2000.

[24] Breiman, L., “Random Forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[25] Chen, T., Guestrin, C., “XGBoost: A Scalable Tree Boosting System,” in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, pp. 785–794, 2016.

[26] Friedman, J.H., “Greedy function approximation: A gradient boosting machine,” *Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.

[27] Langmuir, I., “The adsorption of gases on plane surfaces of glass, mica and platinum,” *J. Am. Chem. Soc.*, vol. 40, no. 9, pp. 1361–1403, 1918.

[28] Temkin, M.I., Pyzhev, V., “Kinetics of ammonia synthesis on promoted iron catalysts,” *Acta Physicochimica URSS*, vol. 12, pp. 327–356, 1940.

## Acknowledgements
The authors acknowledge the support and research facilities provided by the affiliated institution. The Authors are thankful to Dr. Nitin Dongarwar, Prof. Department of Botany, Rashtrasant Tukadoji Maharaj Nagpur University, Nagpur, and Maharashtra, India for his sincere assistance for botanical identification and authentication of plant material employed during present study, giving voucher specimen number 2026/559. Special acknowledgements are due to the faculty members and the research guides for their valuable inputs, constant support.

## Author contributions
Conceptualization: Swaroop laxmi Mudliar. Formal analysis: sayali kawalkar. Funding: No external funding. Software: Dhanashree kuthe., ashish kamble., sai shashank mudliar. Supervision: swarooplaxmi mudliar. Writing-review and editing: swarooplaxmi mudliar., sayali kawalkar., sai shashank mudliar. All author have read and agreed to the published version of the manuscript.

## Competing interests
The authors declare no competing interests.

## Funding
This research received no external funding. This study has not received any specific grant, financial support, or funding from any government, commercial, or non-profit organisation.

## Data Availability
The datasets used and/or analysed during the current study are available from the corresponding author upon reasonable request.

## Additional Information:
The Authentication certificate from RTMNU, Nagpur is attached.