**Text S1. Model development**

The classification features were transformed into numerical features through target encoding to ensure that the literal variables meet the input requirements of the subsequent regression algorithm in modeling (Wan et al., 2022). Meanwhile, to address variations in units and data ranges among variables, standardization was achieved by scaling the raw data using the following equation:

(S1)

Where xi represents the value of the ith input feature; denotes the normalized value of the initial xi with μ being its mean value of xi and σ representing its standard deviation.

The objective of optimizing hyperparameters was to identify the optimal value that yielded the best validation result (Andonie, 2019). The formula for Bayesian optimization is as follows:

(S2)

Where y is the hyperparameter set and belongs to the domain space Y; f(y) is the performance of the model on a validation dataset of hyperparameter x; and y\* is the optimal hyperparameter after optimization.

**Text S2. Experimental validation of the models**

Pig manure (PM) and wheat straw (WS) were obtained from livestock farms and farmlands located at the Base of Sichuan Agricultural University. Wheat straw biochar (WSB) was prepared in the laboratory by pyrolyzing at 600°C for 2 hours under an N2 atmosphere. WS was shredded and mixed uniformly with PM in a 1:2 ratio (on a dry weight basis of PM) with a C/N ratio of approximately 25. The physicochemical properties of the materials are presented in Table S3. The treatment involving the amendment of 8% and 15% WSB was designated as T1 and T2, respectively, while the control without any additives was labeled CK. The experiments were carried out in a 25-liter vessel for 51 days. The moisture content was adjusted to approximately 60.0%. The temperature of the compost pile was measured daily. Samples were collected on days 0, 4, 9, 16, 23, 37, and 51.

The humidity was determined by subjecting the samples to a drying process in an oven at 105°C for approximately 6 hours, and the percentage of weight loss was subsequently recorded. The pH was determined by preparing a 1:10 aqueous extract (w/v) of the fresh compost sample. The sample was extracted with deionized water using an oscillating shaker for 2 hours (Waqas et al., 2018). The temperature in the reactor is monitored daily by a temperature sensor (COS-03-05, Renke, Shandong).

The experiment incorporated four internal maturity indicators, namely germination index (GI), C/N ratio, NO3--N, and NH4+-N, as well as external maturity indicators: CO2, CH4, NH3, and N2O. The GI detection method was based on a seed germination experiment (Zhong et al., 2021). The formula is as follows:

(S3)

Where extract is the extraction solution culture treatment of organic fertilizer, and control is the water culture treatment.

The C/N ratio was determined by measuring the total carbon (TC) and total nitrogen (TN) contents using a Vario EL cube (Wan et al., 2022). The concentrations of NO3--N and NH4+-N were measured using 2 M KCl solution at a mass-to-volume ratio of 1:10, followed by analysis of phenol-disulfonic acid colorimetry methods (Li et al., 2022). The flow rates of N2O, CO2, and CH4 were collected from the reactor in gas sample bags for subsequent quantitative analysis using a gas chromatograph (SP-3510, Beifen, China). NH3 emissions were evaluated and detected through boric acid absorption followed by titration (Mao et al., 2018; Yang et al., 2019). Data were collected in triplicate and presented as the mean ± standard deviation (n = 3), with statistical analysis conducted using IBM SPSS version 26.0 software.

**Text S3. Cosine similarity**

Cosine similarity is a method used to quantify the similarity between two vectors or datasets by calculating the cosine of the angle formed between them (Sun et al., 2020). It is commonly employed in evaluating the similarity between query conditions and records within a dataset, aiding in determining data points belonging to AD. In this context, given two vectors D and 𝐷′, the cosine similarity can be calculated using the following equation (Hanifi et al., 2022):

(S4)

Where 𝜃 represents the angle between D and 𝐷′ in an n-dimensional space, and Di and 𝐷′i represent the ith component of datasets D and 𝐷′ respectively. This method disregards document length; thus, it can effectively handle input data with varying lengths but identical elements (Hanifi et al., 2022).

The model applicability domain (AD) defines the conditions under which a machine learning model's predictions are considered reliable. This involves evaluating whether the input data aligns with the model's training data and if the model can accurately predict outcomes for those inputs. To assess this, we employed cosine similarity to measure the similarity between individual records (converted to vectors).

To establish a suitable similarity threshold, we compared the similarities between records in the training dataset (reference set) and those in the test dataset (exploration set). We computed similarity scores (ranging from 0 to 1) for each test record by identifying its highest similarity to any record in the training set. To determine the appropriate threshold for AD, we incrementally increased the cosine similarity threshold from 0 to 1, analyzing the proportion of test records exceeding or falling below each threshold. For example, with a threshold of 0.7, only test records with a maximum similarity > 0.7 to the training set were retained, while those with lower similarity (< 0.7) were discarded.

Next, we iteratively removed low-similarity records (i.e., those outside the AD) from the test set using a cosine similarity gradient of 0.01. We then evaluated the R² and RMSE values for both the discarded data and the kept data within the AD. Our approach was to discard data with lower model performance while keeping data with higher performance, thereby using their corresponding cosine similarity as a criterion for defining the model's AD.

**Text S4. Indicator hierarchical ranking based on the K-Means algorithm**

The K-means algorithm groups objects into K clusters based on their attributes, iteratively updating the cluster centroids to minimize the distance between each data point and its respective centroid (Velmurugan & Santhanam, 2010). Using the K-means unsupervised classification algorithm, the internal maturity indicators were hierarchically ranked through the following steps:

(1) Each pair of the four internal maturity indicators was input into the K-means clustering algorithm (k=2). Two clusters were used because the purpose was to classify the data into Cluster 0 (mature) and Cluster 1 (immature) compost. Pairing two indicators at a time, rather than using all four, allowed us to analyze more data, as only 148 records contained simultaneous reports of all four indicators.

(2) To verify the accuracy of the K-means clustering results, we compared the predicted maturity/immaturity clusters to the reported compost maturity statuses, which were classified based on the predefined statistical maturity thresholds.

(3) Since each K-means process involved two indicators, we shuffled one of them and analyzed how the clustering results changed compared to the original data without shuffling. Greater changes in clustering accuracy indicated a higher importance of the shuffled indicator in determining maturity. Having already verified the accuracy of the K-means clustering in Step 2, we assessed the percentage change in clustering accuracy. A larger change in accuracy suggested that the shuffled indicator had a more significant influence on distinguishing between immature and mature compost, thus identifying its importance in assessing compost maturity.

For instance, when comparing the C/N ratio (as X) with GI (as Y), the accuracy of k-means was evaluated by comparing how many reported maturity indicators were correctly clustered (according to the statistical maturity threshold, cluster 0 was the mature cluster). The formulas are as follows:

(S5-1)

(S5-2)

(S6-1)

(S6-2)

(S7-1)

(S7-2)

Unsupervised clustering was performed using the input of C/N ratio and GI. Then we randomly shuffled the C/N ratio data, and randomly shuffled GI data separately, followed by re-clustering. By comparing the relative changes in distribution accuracy between the original data (eq. S5), randomly shuffled C/N ratio data (eq. S6), and randomly shuffled GI data (eq. S7) within the mature cluster, we can determine the hierarchical importance of the C/N ratio and GI. For instance, when the C/N ratio values were shuffled (Figure S10A), the k-means accuracy decreased by 12.26% (from 56.13% to 43.87%). In comparison, shuffling the GI values resulted in a larger accuracy drop of 15.48% (from 56.13% to 40.65%). This greater absolute change in accuracy indicates that GI holds relatively higher importance than the C/N ratio.

Moreover, across all binary k-means results (Figure S10), the classification accuracy of the mature GI group (i.e., GI > 83.34% in the original data) within the mature cluster remained stable at 100%, regardless of the random shuffling of other maturity indicators. This suggests that data identified as mature by GI (> 83.34%) can be reliably predicted by all four maturity indicators, highlighting GI’s overriding importance. In scenarios where we aim to classify compost maturity, it can be difficult to establish clear cutoffs for maturity due to multiple indicators. However, the stability of the GI-based classification, regardless of the values of other indicators, suggests that GI can serve as a decisive factor. When GI exceeds 83.34%, the composting process should be considered mature, underscoring GI's dominant role among the four indicators. This is further confirmed by the results of binary k-means comparisons.

The reliability of the results obtained from binary comparison was validated by a 3D K-means clustering incorporating a four-element comparison. Specifically, this iterative process involved aggregating data from the four internal maturity indicators into a set of four metadata for unsupervised classification using the K-means clustering algorithm. When simultaneously clustering the four indicators, the number of records was limited to 148 due to four sets of data representing a single cluster point. Similarly, both the original data and randomly shuffled C/N ratio, GI, NH4+-N, and NO3--N data were subjected to clustering analysis. The relative influence of different shuffled data and original data clustering results was investigated respectively, and the relative importance hierarchy of the indicators was finally determined.

**Text S5. Weight allocation of compost maturity indicators based on analytic hierarchy process (AHP)**

The AHP is a widely adopted systematic approach for assessing the relative importance of different indicators in addressing complex multi-objective decision problems (Zhang et al., 2024). An importance evaluation matrix was established using the analytic hierarchy process (AHP), and weights were assigned to the four internal maturity indicators (Table S5). We compared two elements in constructing the evaluation matrix on an importance scale:

1= “Equally important” (S8-1)

3= “Moderately important” (S8-2)

5= “Strongly important” (S8-3)

7= “Very strongly important” (S8-4)

9= “Extremely important” (S8-5)

2, 4, 6, 8= “Intermediate values” (S8-6)

In the given formula, scales 1-9 correspond to different values; when comparing the latter with the former, the reciprocal is used as a scaling factor (Zhang et al., 2024). The eigenmatrix A and indicator relative weight can be calculated using the following equations:

(S9)

(S10)

Where represents the result of comparing the importance between factor i and factor j, while n denotes the dimensionality of the matrix. Furthermore, the consistency assessment for evaluating the AHP importance matrix is achieved using the subsequent formula:

(S11)

(S12)

Where C.I. is the consistency index; R.I. is the average random consistency index, which is determined by the order of the importance evaluation matrix. The test coefficient (C.R.) is used to evaluate whether the importance evaluation matrix passes the conformance test (Zhong et al., 2022). When the C.R. value ≤ 0.1, the conformance test is passed according to the AHP standard (Zhang et al., 2024). If this condition is not met, adjustments must be made to ensure the accuracy of the importance evaluation matrix.

**Text S6. Basic formula of genetic algorithm**

The genetic algorithm (GA) is an optimization method grounded in biological principles, extensively employed across diverse domains to generate high-quality optimization schemes by means of selection, crossover, and mutation operations (Akbarifard et al., 2024). The compost maturity day was estimated in this study by employing a GA. The initial combination of different process parameters was randomly generated based on the input data's initial values of process parameters (Day\_P, pH\_P, TEMP\_P, and MC\_P). Subsequently, the process parameters were continuously iterated and utilized as input features in the model to ensure convergence of the predicted value for the current maturity indicator towards the statistical maturity thresholds. The relatively better combination of process parameters was then output, followed by random selection for cross-transform processing and random change processing. By randomly modifying certain parameter values within each optimal combination, this introduced new combinations and increased the chances of finding better solutions, while enhancing diversity and avoiding local optima. Finally, we outputted the best process parameters from all four internal maturity indicators and weighted the regulated Day\_P for each internal maturity indicator. This process was repeated 15 times to obtain an average estimate of the final assessment maturity day. The formulas of GA are as follows:

ML models：C/N ratio, GI, NH4+-N, NO3--N= ML(Inputs) (S13)

GA optimization: C/N ratio prediction ≤ C/N ratio Statistical maturity threshold; NH4+-N prediction ≤ NH4+-N Statistical maturity threshold;

GI prediction ≥ GI Statistical maturity threshold; NO3--N prediction ≥ NO3--N Statistical maturity threshold (S14)

The ML(Inputs) in the given expression represents the optimal predictive models that must be established and optimized for predicting compost maturity indicators.

**Text S7. Visualization of regulation and prediction for engineering applications**

The prediction of compost maturity indicators has been implemented in Python, along with the estimation of maturity day using GA, to establish a web-based intelligent prediction framework for biochar amendment composting. The following six steps were undertaken: (1) The developed multiple models were exported in a "pickle" format; (2) Web interfaces were developed using HTML front-end technologies to expedite development; (3) A back-end application was constructed utilizing Python's Web framework (Flask), which handled web requests and invoked models for predictions; (4) API interfaces were written to receive data from the front end and transmit it to the weight allocation model with unified maturity indicator for prediction; (5) In the back-end code, the optimal weights were imported and loaded, followed by writing code to perform prediction operations; and (6) The weight allocation model was invoked based on data sent by the front end, predictions were made, and results were returned.

**References**

Akbarifard, S., Madadi, M.R., Zounemat-Kermani, M. 2024. An artificial intelligence-based model for optimal conjunctive operation of surface and groundwater resources. Nat. Commun., 15(1), 553. <https://dx.doi.org/https://doi.org/10.1038/s41467-024-44758-6>.

Andonie, R. 2019. Hyperparameter optimization in learning systems. J. Membr. Comput., 1(4), 279-291. <https://dx.doi.org/https://doi.org/10.1007/s41965-019-00023-0>.

Hanifi, M., Chibane, H., Houssin, R., Cavallucci, D. 2022. Problem formulation in inventive design using Doc2vec and cosine similarity as artificial intelligence methods and scientific papers. Eng. Appl. Artif. Intell., 109, 104661. <https://dx.doi.org/https://doi.org/10.1016/j.engappai.2022.104661>.

Li, Y., Zhang, G., Xu, P., Zhou, S., Li, Y., Ma, L., Yang, Z., Wu, Y. 2022. Effects of Exogenous Bacterial Agents on Material Transformation and Microbial Community Composition during Composting of Tomato Stalks. Sustainability, 14(23), 16284. <https://dx.doi.org/https://doi.org/10.3390/su142316284>.

Mao, H., Lv, Z., Sun, H., Li, R., Zhai, B., Wang, Z., Awasthi, M.K., Wang, Q., Zhou, L. 2018. Improvement of biochar and bacterial powder addition on gaseous emission and bacterial community in pig manure compost. Bioresour. Technol., 258, 195-202. <https://dx.doi.org/https://doi.org/10.1016/j.biortech.2018.02.082>.

Sun, X., Zhang, X., Muir, D.C.G., Zeng, E.Y. 2020. Identification of potential PBT/POP-like chemicals by a deep Learning approach based on 2D structural features. Environ. Sci. Technol., 54(13), 8221-8231. <https://dx.doi.org/https://doi.org/10.1021/acs.est.0c01437>.

Velmurugan, T., Santhanam, T. 2010. Computational complexity between K-means and K-medoids clustering algorithms for normal and uniform distributions of data points. J. Computer Sci., 6(3), 363. <https://dx.doi.org/https://doi.org/10.3844/jcssp.2010.363.368>.

Wan, X., Li, J., Xie, L., Wei, Z., Wu, J., Tong, Y.W., Wang, X., He, Y., Zhang, J. 2022. Machine learning framework for intelligent prediction of compost maturity towards automation of food waste composting system. Bioresour. Technol., 365, 128107. <https://dx.doi.org/https://doi.org/10.1016/j.biortech.2022.128107>.

Waqas, M., Nizami, A., Aburiazaiza, A., Barakat, M., Ismail, I., Rashid, M. 2018. Optimization of food waste compost with the use of biochar. J. Environ. Manage., 216, 70-81. <https://dx.doi.org/https://doi.org/10.1016/j.jenvman.2017.06.015>.

Yang, F., Li, Y., Han, Y., Qian, W., Li, G., Luo, W. 2019. Performance of mature compost to control gaseous emissions in kitchen waste composting. Sci. Total Environ., 657, 262-269. <https://dx.doi.org/https://doi.org/10.1016/j.scitotenv.2018.12.030>.

Zhang, Y., Wen, Z., Zhang, J., Huang, J., Qi, J., Li, M. 2024. Study on stability evaluation of goaf based on AHP and EWM—taking the northern new district of Liaoyuan city as an example. Sci. Rep., 14(1), 17876. <https://dx.doi.org/https://doi.org/10.1038/s41598-024-68858-x>.

Zhong, B., An, X., An, W., Xiao, X., Li, H., Xia, X., Zhang, Q. 2021. Effect of bioaugmentation on lignocellulose degradation and antibiotic resistance genes removal during biogas residues composting. Bioresour. Technol., 340, 125742. <https://dx.doi.org/https://doi.org/10.1016/j.biortech.2021.125742>.

Zhong, C., Yang, Q., Liang, J., Ma, H. 2022. Fuzzy comprehensive evaluation with AHP and entropy methods and health risk assessment of groundwater in Yinchuan Basin, northwest China. Environ. Res., 204, 111956. <https://dx.doi.org/https://doi.org/10.1016/j.envres.2021.111956>.