Journal of Computing and Electronic Information Management
ISSN: 2413-1660 | Vol. 20, No. 3, 2026

A Happiness Prediction Model Based on Performance-Diversity Dual-Constraint Dynamic Weighted Ensemble

Siqi Zhang, Jing Zhang

School of Computer Science and Technology, Taiyuan Normal University, Jinzhong 030619, China

Abstract: To address the problems of insufficient generalization ability of traditional single machine learning models in happiness prediction tasks and the failure of fixed-weight ensemble models to adapt to sample-level prediction differences, this paper proposes a Performance-Diversity Dual-constraint Dynamic Weighting Ensemble (PD-DWE) model. First, a diversified base model pool including Random Forest, K-Nearest Neighbors, Support Vector Machine, Gradient Boosting Tree, XGBoost, Multilayer Perceptron, and LightGBM is constructed, and hyperparameter optimization of the base models is completed through a two-stage random search. Second, based on the 10-fold stratified cross-validation framework, high-quality model selection is performed using an F1-score threshold in each fold, and diversity constraint is achieved by combining predictive correlation filtering. Finally, a dynamic weight calculation strategy integrating the global performance of models and sample-level prediction confidence is designed, and the optimal classification threshold is adaptively searched to complete the final prediction. Experimental results show that the proposed PD-DWE model achieves an F1-score of 79.1% on the test set, representing a 12.43% relative improvement compared with the optimal single model, and an AUC value of 69.61%. Its comprehensive prediction performance is significantly superior to that of single machine learning models, providing a more stable and efficient technical solution for the quantitative prediction of happiness.

Keywords: Happiness Prediction; Ensemble Learning; Dynamic Weighting; Machine Learning; Model Fusion.

1. Introduction

Residents h appiness is a core indicator for measuring the quality of social development and the level of peoples livelihood security. Its quantitative prediction and analysis of influencing factors have become a research hotspot in the interdisciplinary field of sociology, economics, and computer science. Traditional happiness research mostly relies on descriptive statistics and linear regression models, which are difficult to capture the nonlinear correlations between happiness and influencing factors, resulting in significant limitations in prediction accuracy and generalization ability.

With the development of machine learning technology, nonlinear models such as Decision Tree, Random Forest, and Gradient Boosting Tree have been widely applied to happiness prediction tasks. Compared with traditional statistical methods, their ability to fit complex nonlinear relationships has been significantly improved[1]. However, existing studies mostly adopt a single machine learning model. Limited by data distribution and the models own assumptions, a single model is prone to overfitting and insufficient generalization ability, and its prediction stability is poor on small-sample and imbalanced datasets.

Ensemble Learning, by integrating the prediction results of multiple base models, can effectively reduce the bias and variance of a single model, and is a mainstream method to improve the performance of classification tasks[2]. Most existing ensemble models adopt a fixed-weight fusion strategy, that is, assigning fixed weights to each base model through manual rules, which cannot adapt to the differences in the prediction ability of each base model under different samples—some base models have extremely high prediction accuracy on samples with specific feature distributions but perform poorly on other samples, and fixed weights cannot give full play to the local advantages of base models.

To address the above issues, this paper takes the binary classification prediction of residents happiness as the research scenario and proposes the PD-DWE dynamic weighted ensemble model. The main contributions of this paper are as follows:

1. Construct a diversified base model pool containing 7 classic machine learning models, complete hyperparameter optimization through a two-stage random search, and ensure the complementarity between base models by combining the differentiated design of model structures and fitting logics;
2. Design a Performance-Diversity dual-constraint base model selection mechanism, filter out low-performance and highly homogeneous base models in each fold of cross-validation, and avoid interference from invalid models on the ensemble results;
3. Propose a dynamic weight calculation strategy integrating global performance and sample-level confidence, assign exclusive fusion weights to each sample, and adaptively search for the optimal classification threshold to maximize the F1 performance of the ensemble model;
4. Complete model training and evaluation through 10-fold stratified cross-validation, compare the prediction performance between single models and the proposed PD-DWE model, and verify the effectiveness and stability of the proposed model.

2. Related Work

2.1. Research Status of Happiness Prediction

Research on happiness prediction can be divided into two stages: traditional statistical methods [3]and machine learning methods[4]. Early studies were mostly based on questionnaire survey data, adopting statistical methods such as multiple linear regression and structural equation modeling to analyze the impacts of factors including income, education and health on happiness. Such methods feature strong interpretability, yet they are deficient in the ability to fit nonlinear correlations and yield relatively low prediction accuracy.

In recent years, machine learning methods[5] have become the mainstream technology for happiness prediction. Relevant studies have demonstrated that ensemble models such as Random Forest and Gradient Boosting Tree achieve significantly better performance than linear models in happiness prediction tasks; some studies have realized the multi-classification prediction of residents happiness levels via models like Support Vector Machine[6] and neural networks, which verifies the applicability of machine learning methods in this scenario. However, existing studies mostly focus on hyperparameter tuning and feature engineering for single models, with inadequate research on the optimization of multi-model fusion, thus failing to fully exploit the performance potential of ensemble learning.

2.2. Research Progress in Ensemble Learning

The core idea of ensemble learning is to integrate the prediction results of multiple base learners through a "divide and conquer" strategy, thereby reducing the generalization error of a single model. Mainstream ensemble frameworks can be divided into three categories: Bagging, Boosting and Stacking. Bagging trains multiple independent base models by random sampling of the dataset and achieves fusion through a voting mechanism, with the Random Forest as its representative model; Boosting conducts iterative training that focuses on samples mispredicted by previous models, and its typical models include AdaBoost, XGBoost and LightGBM; Stacking learns the fusion rules of base models via a meta-learner, enabling more flexible model fusion.

Weighted voting is the most commonly used fusion strategy in ensemble learning[7]. Most existing studies adopt fixed weight assignment, such as allocating weights based on the accuracy and F1-score of base models. Such methods feature simple calculation yet fail to adapt to sample-level differences in prediction performance. Dynamic weighted ensemble, by contrast, assigns exclusive weights to each test sample. Current research on this strategy mostly designs weights based on prediction confidence and fuzzy membership, and it has achieved better performance than fixed weighting in scenarios like medical diagnosis and financial risk control. However, its application in happiness prediction remains scarce, and it lacks dual constraints on the performance and diversity of base models.

3. PD-DWE Happiness Prediction Model

3.1. Overall Framework of the PD-DWE Model

The PD-DWE happiness prediction model proposed in this paper is constructed based on the 10-fold stratified cross-validation framework, as illustrated in Fig. 1. Its overall process consists of four core stages:

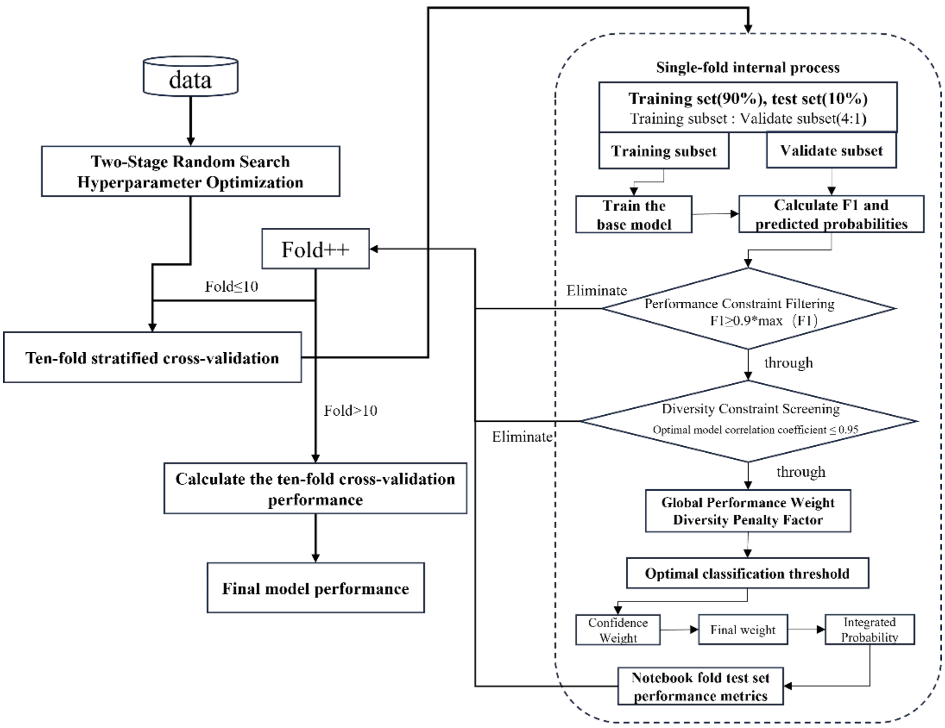

Fig. 1 Model Flow Diagram

(1) Construction of a diversified base model pool: A complementary base model pool is built on the basis of the seven aforementioned machine learning models, through differentiated parameter configuration and two-stage hyperparameter optimization.

(2) Performance-Diversity dual-constraint base model selection: In each fold of cross-validation, high-quality model selection is completed based on the validation set to filter out low-performance and highly homogeneous invalid models, retaining base models with both satisfactory performance and good complementarity.

(3) Sample-level dynamic weight calculation: The global performance of models is integrated with sample-level prediction confidence to assign exclusive fusion weights to each test sample, achieving sample-level adaptive weighting.

(4) Adaptive classification threshold optimization: The optimal classification threshold that maximizes the F1-score is searched for on the validation set to accomplish the final sample classification prediction and performance evaluation.

core process as follows:

1. Performance-Constrained Selection

The F1-score of each base model is calculated on the validation subset. Taking the F1-score of the optimal base model in the current fold as the benchmark, only models with an F1-score no less than 90% of the benchmark value are retained, while low-performance invalid models are filtered out. This ensures that the selected models possess qualified basic predictive ability.

2. Diversity-Constrained Selection

Taking the optimal-performing base model in the current fold as the reference, the Pearson correlation coefficients[9] of the predicted probabilities between the remaining candidate models and the optimal model on the validation set are calculated. Highly homogeneous models with a correlation coefficient greater than 0.95 are filtered out, and only base models with strong complementarity in prediction results are retained, thus avoiding performance loss caused by "homogenization averaging" during the ensemble process.

Through dual-constraint selection, the finally selected base models not only guarantee satisfactory basic predictive performance but also have sufficient complementarity in prediction results, laying a solid foundation for subsequent dynamic weighted fusion.

3.2. Construction of Base Model Pool and Two-Stage Hyperparameter Optimization

To ensure the complementarity among base models, seven classic machine learning models with differentiated structures and complementary fitting logics are selected in this paper to form the base model pool, including K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Random Forest (RF), Gradient Boosting Tree (GradientBoost), XGBoost, Multilayer Perceptron (MLP), and LightGBM. These models cover four categories, namely tree ensemble, nearest neighbor, kernel method and neural network, and are capable of capturing data features from different perspectives.

To fully exploit the predictive potential of each base model, a two-stage random search[8] strategy is adopted for hyperparameter optimization, with the F1-score as the optimization objective and the whole evaluation process based on 5-fold cross-validation. The first stage is wide-range rough search: a wide search space covering the core hyperparameters is set for each base model, and the approximate range of the optimal parameters is quickly located through 15 random iterations. Meanwhile, parameter validity constraints (e.g., min_samples_split≥2, subsample∈[0,1]) are imposed to ensure the rationality of the search results. The second stage is narrow-range fine search: a narrow search space is constructed by scaling the optimal parameters obtained from the rough search by 50% up and down around their values, and refined tuning is then completed through 10 random iterations. Finally, the optimal hyperparameter combination for each base model is derived.

3.3. 10-Fold Stratified Cross-Validation Experimental Framework

To ensure the objectivity and generalization ability of model evaluation and completely avoid data leakage, this paper adopts a 10-fold stratified cross-validation framework for model training and evaluation, with the core process as follows:

The preprocessed happiness dataset is randomly divided into 10 mutually exclusive subsets with consistent label distribution, ensuring that the ratio of positive and negative samples in each fold is the same as that in the entire dataset; nine subsets are selected as the training set each time, and the remaining one is used as the test set. This process is executed iteratively for 10 times to guarantee that each subset serves as the test set exactly once. For the training set in each iteration, it is further split into a training subset and a validation subset at a ratio of 4:1: the training subset is exclusively used for the training of base models, while the validation subset is only applied to base model selection, weight calculation and threshold optimization, and is not involved in model training throughout the process. In each iteration, model training and parameter optimization are completed solely with the training data of the current fold, and the test set is only used for the final performance evaluation. This approach completely eliminates data leakage and ensures the reliability and reproducibility of the experimental results.

3.4. Performance-Diversity Dual-Constraint Base Model Selection Mechanism

To avoid the interference of low-performance and highly homogeneous models on ensemble results, this paper conducts dual-constraint base model selection based on the validation subset in each fold of cross-validation, with the core process as follows:

1. Performance-Constrained Selection

The F1-score of each base model is calculated on the validation subset. Taking the F1-score of the optimal base model in the current fold as the benchmark, only models with an F1-score no less than 90% of the benchmark value are retained, while low-performance invalid models are filtered out. This ensures that the selected models possess qualified basic predictive ability.

2. Diversity-Constrained Selection

Taking the optimal-performing base model in the current fold as the reference, the Pearson correlation coefficients[9] of the predicted probabilities between the remaining candidate models and the optimal model on the validation set are calculated. Highly homogeneous models with a correlation coefficient greater than 0.95 are filtered out, and only base models with strong complementarity in prediction results are retained, thus avoiding performance loss caused by "homogenization averaging" during the ensemble process.

Through dual-constraint selection, the finally selected base models not only guarantee satisfactory basic predictive performance but also have sufficient complementarity in prediction results, laying a solid foundation for subsequent dynamic weighted fusion.

3.5. Sample-level Dynamic Weight Calculation Strategy

This paper designs a dynamic weight calculation strategy that integrates the global performance of models and sample-level prediction confidence, assigning exclusive fusion weights to each test sample. The calculation process consists of five steps:

1.Global Performance Base Weight: The square of the F1-score of a base model on the validation set is used as the base weight, with the formula $w_i^{global} = (F1_i)^2$. This amplifies the weight advantage of high-performance models through the squaring operation.

2.Diversity Penalty Factor: Calculate the average predicted correlation coefficient of the i-th base model with the remaining selected models as $\rho_i = \frac{1}{n-1}\sum_{j\neq 1}\rho_{ij}$, where ρ_{ij} denotes the Pearson correlation coefficient of the predicted probabilities between model i and model j on the validation set. The diversity penalty factor is defined as $d_i = 1-\rho_i$, which enables models with low correlation with other models to obtain higher weights.

3.Sample-level Confidence Weight: The models confidence in the current test sample is measured based on prediction entropy, with lower entropy indicating higher confidence. The calculation formula for binary classification prediction entropy is $e_{ik} = -(p_{ik}\log(p_{ik}+\epsilon)+(1-p_{ik})\log(1-p_{ik}+\epsilon))$, where p_{ik} is the positive class prediction probability of the i-th model for the k-th sample, and $\epsilon = 10^{-8}$ is set to avoid logarithmic calculation errors. The confidence weight is defined as $c_{ik} = 1 - e_{ik} / \log 2$, which is the complement of the normalized entropy.

4.Final Dynamic Weight Normalization: The three aforementioned indicators are integrated to derive the final weight of the i-th model for the k-th sample as $w_{ik} = w_i^{global} \cdot d_i \cdot c_{ik}$, and a normalization process is performed to ensure the sum of weights equals 1.

5.Dynamic Weighted Probability Fusion: Based on the normalized dynamic weights, a weighted summation of the predicted probabilities of the base models is conducted to obtain the ensemble positive class prediction probability of the test sample as $P_k = \sum_i w_{ik} \cdot p_{ik}$.

3.6. Adaptive Classification Threshold Optimization

Traditional binary classification tasks default to 0.5 as the classification threshold. However, in scenarios with imbalanced datasets or uneven model output distributions, 0.5 is not the optimal threshold. In this paper, with the maximization of the F1-score as the optimization objective, an exhaustive search for the optimal classification threshold T is conducted on the validation subset within the interval [0.1,0.9] with a step size of 0.01. The final classification of samples is then determined based on this optimal threshold:

$$\hat{y}_k = \begin{cases} 1, & P_k \geq T \\ 0, & P_k < T \end{cases} \tag{1}$$

where $\hat{y}_k$ denotes the final classification result of the k-th test sample, 1 represents high happiness, and 0 represents low happiness.

4. Experimental Results and Analysis

4.1. Dataset and Experimental Setup

In this paper, the public Somerville Happiness Dataset[10] from the UCI Machine Learning Repository is selected to validate the proposed method. This dataset consists of 143 samples and 6 features, with the variable definitions and statistical characteristics presented in Table 1, and the visualization of inter-feature correlations shown in Fig.2.

The experiments were implemented based on Python 3.11, with core algorithm libraries including scikit-learn, xgboost, lightgbm, numpy and pandas, and visualization realized via matplotlib. The random seed was fixed at 42 for all experiments to ensure the reproducibility of results;

Table 1. Statistical Description of the Somerville Happiness Survey Dataset

Features	Symbol	Mean	Std	Min	25%	50%	75%	Max
The availability of information about the city services	X1	4.31	0.80	1.0	4.0	5.0	5.0	5.0
The cost of housing	X2	2.54	1.12	1.0	2.0	3.0	3.0	5.0
The overall quality of public schools	X3	3.27	0.99	1.0	3.0	3.0	4.0	5.0
Your trust in the local police	X4	3.70	0.89	1.0	3.0	4.0	4.0	5.0
The maintenance of streets and sidewalks	X5	3.62	1.31	1.0	3.0	4.0	4.0	5.0
The availability of social community events	X6	4.22	0.85	1.0	4.0	4.0	5.0	5.0
Decision attribute (D) with values 0 (unhappy) and 1 (happy)	D	0.54	0.50	0.0	0.0	1.0	1.0	1.0

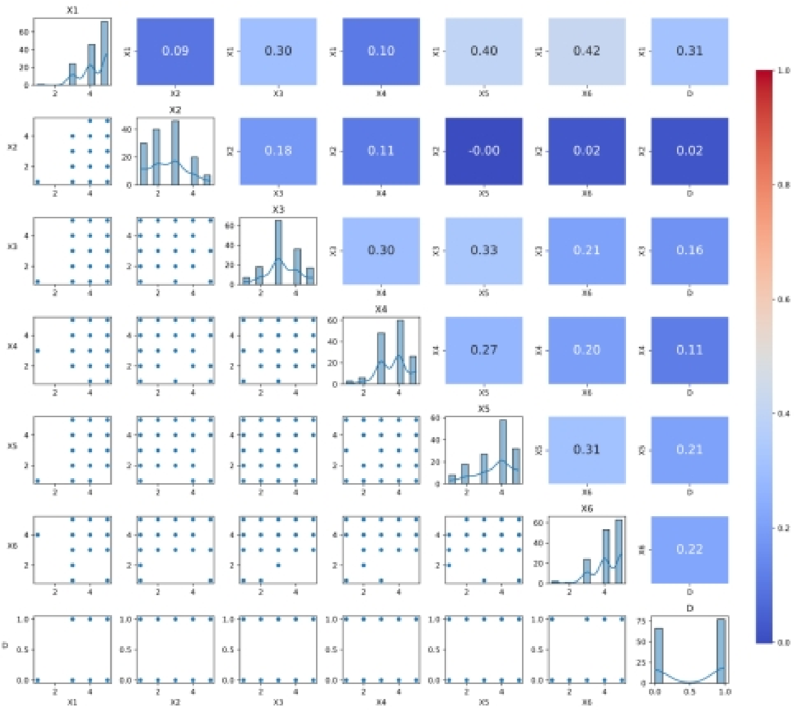

Fig. 2 Pairwise relationships between features

4.2. Evaluation Metrics

To comprehensively evaluate the performance of the happiness prediction model, this study adopts multiple evaluation metrics: Accuracy, Precision, Recall, F1-score, and AUC value. The evaluation metrics of the model can be calculated according to the following Equations (2) to (6):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5)$$

$$AUC = \frac{\sum_{i=1}^{M} R_i - \frac{M(M+1)}{2}}{M \times N} \quad (6)$$

where TP (True Positive) denotes true positives, TN (True Negative) denotes true negatives, FP (False Positive) denotes false positives, FN (False Negative) denotes false negatives, R_i is the rank position of the i-th positive sample after sorting (the rank position starts from 1), M is the number of positive samples, and N is the number of negative samples.

4.3. Model Performance Comparison

Table 2 shows the average test set performance and overall ranking of seven basic machine learning models and the proposed PD-DWEM ensemble model under 10-fold cross-validation. Fig. 3 intuitively presents the performance bar chart of each model. From the experimental results of the base models, KNN achieves the best overall performance among all basic models, with an F1-score of 0.6667, and its accuracy and AUC value are the highest among the base models. SVM shows the most outstanding recall performance among the base models, reaching 0.7143, and has the optimal recognition coverage for positive samples of high happiness. Models such as LightGBM and RF also present respective advantages in precision and indicator balance. In contrast, MLP exhibits the weakest overall performance among all base models and will be filtered out in the subsequent dual-constraint selection stage to avoid interference with the ensemble effect.

Table 2. Performance Comparison between PD-DWEM and Other Models

Algorithm	Accuracy	Precision	Recall	F1	AUC	Rank
KNN	0.6364	0.6582	0.6753	0.6667	0.6407	2
SVM	0.5874	0.5978	0.7143	0.6509	0.6297	7
RF	0.6154	0.6410	0.6494	0.6452	0.6392	4
XGBoost	0.5944	0.6234	0.6234	0.6234	0.5861	5
LightGBM	0.6154	0.6667	0.5714	0.6154	0.6396	3
GradientBoost	0.5874	0.6216	0.5974	0.6093	0.5890	6
MLP	0.5385	0.5679	0.5974	0.5823	0.5688	8
PD-DWEM	0.7413	0.7000	0.9091	0.7910	0.6961	1

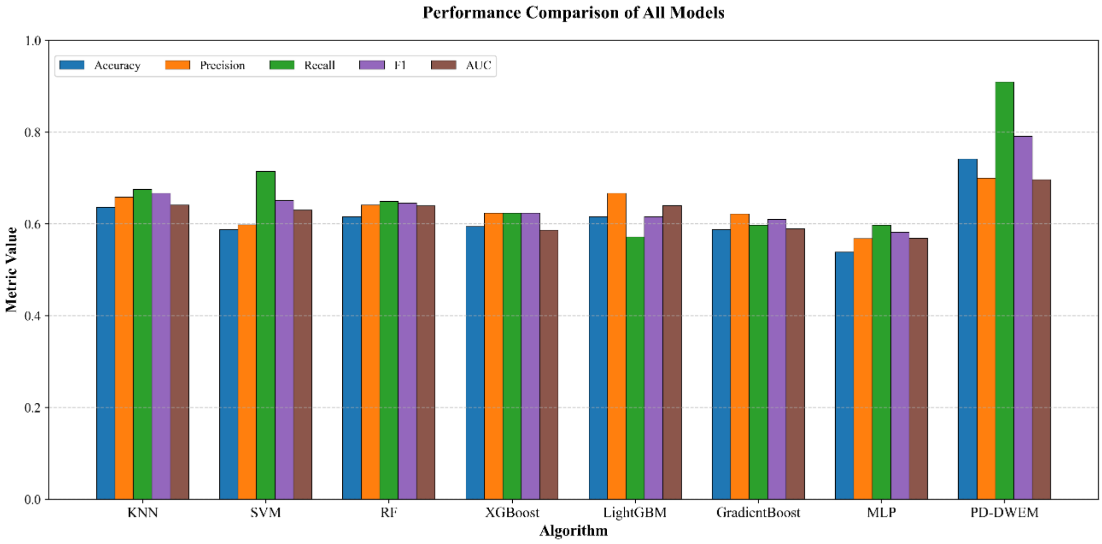

Fig. 3 Bar Chart of Model Performance

The PD-DWEM ensemble model proposed in this paper outperforms all basic models in all core evaluation metrics and ranks first among all models, which fully verifies the effectiveness and advancement of the proposed ensemble framework. In terms of core classification performance, the F1-score of the PD-DWEM model reaches 0.7910, representing a 18.64% relative improvement compared with the best-performing base model KNN, which greatly optimizes the overall performance of the classification task. In classification accuracy, the model achieves an accuracy of 0.7413, with an increase of 16.48% over the optimal base model KNN, representing a qualitative leap in the overall classification correctness. Particularly notably, the recall rate of the PD-DWEM model reaches 0.9091, which not only far exceeds that of SVM (the base model with the highest recall), but also represents a 34.65% improvement compared with the optimal base model KNN. This indicates that the model can greatly reduce the missed detection rate of positive samples with high happiness, and the recognition coverage of target samples has been improved by an order of magnitude, demonstrating strong practical value in real-world application scenarios of happiness prediction. Meanwhile, the precision of the model remains at a high level of 0.7000, and the AUC value reaches 0.6961, an increase of 8.64% compared with the optimal base model. While greatly improving the recall rate, there is no significant drop in precision, achieving a better balance between precision and recall. The ability to distinguish between positive and negative samples is also significantly superior to all basic models.

The above results fully demonstrate that the performance-diversity dual-constraint base model selection mechanism and the sample-level dynamic weighted fusion strategy designed in this paper can effectively leverage the predictive advantages of different base models, avoid the performance shortcomings of single models, and ultimately achieve a comprehensive improvement in happiness prediction performance.

4.4. ROC Curve Analysis

Fig. 4 shows the ROC curves of all base models and the proposed PD-DWE model in this paper, where our model is marked with a red dashed line.It can be observed from the ROC curves that:The ROC curve of the proposed PD-DWE model lies above most base models, and the true positive rate is significantly improved in the low false positive rate region. This indicates that under the premise of controlling the false detection rate, the model's ability to identify positive samples is remarkably superior to single models, showing stronger practical application value.

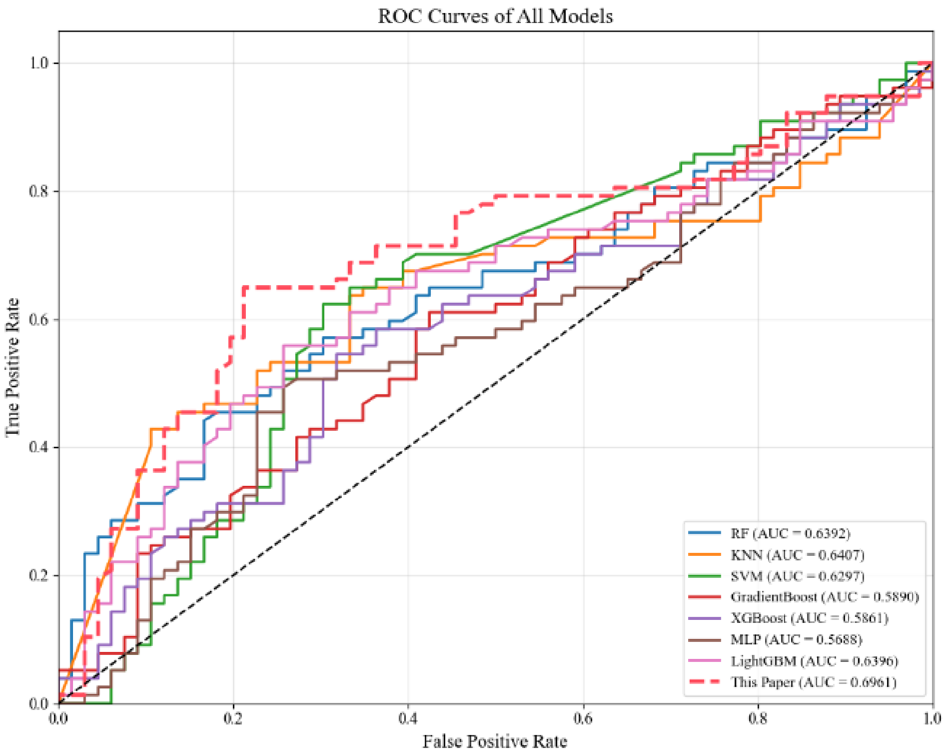

Fig. 4 ROC Curve Comparison of All Models

5. Summary

This paper conducts a systematic study focusing on the practical demands of quantitative prediction for residents' happiness and the application shortcomings of existing machine learning methods in this scenario. Targeting the core pain points in the current happiness prediction field—namely, the insufficient generalization ability of single models and the failure of fixed weights in traditional ensemble models to adapt to sample-level prediction differences, a complete happiness prediction framework, namely Performance-Diversity Dual-Constraint Dynamic Weighted Ensemble (PD-DWE), is constructed.

In the model construction phase, seven classic machine learning models with complementary structures are selected to build a diversified base model pool: K-Nearest Neighbor, Support Vector Machine, Random Forest, Gradient Boosting Decision Tree, XGBoost, Multilayer Perceptron, and LightGBM. A two-stage random search strategy is adopted to optimize the hyperparameters of each base model, laying a solid single-model performance foundation for ensemble fusion. Meanwhile, based on the 10-fold stratified cross-validation framework, a two-layer base model selection mechanism combining performance threshold filtering and prediction correlation constraint is designed to avoid the interference of low-performance and highly homogeneous models on ensemble results from the source. On this basis, a dynamic weight calculation strategy integrating global model performance and sample-level prediction confidence is proposed, coupled with adaptive classification threshold optimization, realizing adaptive prediction for samples with different feature distributions.

The proposed PD-DWE model not only significantly outperforms single base models in core evaluation metrics such as F1-score and recall rate, but also achieves stable performance gains compared with traditional ensemble strategies including equal-weight voting and fixed F1-weight voting. It effectively remedies the deficiencies of existing happiness prediction models in stability and sample adaptability.

This study not only provides an efficient and reproducible technical solution for the accurate quantitative prediction of residents' happiness, but also offers a reusable ensemble learning framework for classification and prediction tasks on small-sample social survey data. It can provide data-level technical support for policy formulation and effect evaluation in the people's livelihood domain. In future work, the prediction performance and practical application value of the model can be further enhanced by expanding feature dimensions, integrating model interpretability analysis tools, and optimizing the structure of the ensemble framework.

References

[1] Zhang, C. Y., Zhu, W. L., Li, X. R., Chen, G. Z., & Peng, M. Y. (2023). The outcome prediction model of acute stroke: Comparison between machine learning and traditional regression model. Chinese Journal of CT and MRI, 21(7), 24-26. https://doi.org/10.3969/j.issn.1672-5131.2023.07.008

[2] Airlangga G , Liu A .A Hybrid Gradient Boosting and Neural Network Model for Predicting Urban Happiness: Integrating Ensemble Learning with Deep Representation for Enhanced Accuracy[J].Machine Learning & Knowledge Extraction, 2025, 7(1).DOI:10.3390/make7010004.

[3] Avc Z , Bardak T , Bardak S ,et al.Evaluation of Employee's Happiness by Data Mining and Statistical Methods in the Process of Internal Public Relations[C]//International Congress on Engineering and Life Science.2018.

[4] Geici E , Yildiz A , Alp S .The Hybrid Best Worst Integrated Machine Learning Methodology to Clustering the Countries Based on the Happiness Index[J].Access, IEEE, 2025, 13(000):98076-98087.DOI:10.1109/ACCESS.2025.3576378.

[5] Alhur A A , Alsahmmari M , Al-Khattab M A .Can Lifestyle Habits Predict Happiness? An Exploratory Machine Learning Study Using a Visual Data Mining Platform[J].Journal of Pioneering Medical Sciences, 2025, 14(6).

DOI:10.47310/jpms2025140607.

[6] Khemraj S, Thepa P C A, Chi H, et al. Prediction of world happiness scenario effective in the period of COVID-19 pandemic, by artificial neuron network (ANN), support vector machine (SVM), and regression tree (RT)[J]. NVEO-NATURAL VOLATILES & ESSENTIAL OILS Journal| NVEO, 2021, 8(4): 13944-13959.

[7] Osamor V C, Okezie A F. Enhancing the weighted voting ensemble algorithm for tuberculosis predictive diagnosis[J]. Scientific Reports, 2021, 11(1): 14806.

[8] Malakouti S M, Menhaj M B, Suratgar A A. Applying grid search, random search, Bayesian optimization, genetic algorithm, and particle swarm optimization to fine-tune the hyperparameters of the ensemble of ML models enhances its predictive accuracy for mud loss[J]. 2024.

[9] Šverko Z, Vrankić M, Vlahinić S, et al. Complex Pearson correlation coefficient for EEG connectivity analysis[J]. Sensors, 2022, 22(4): 1477.

[10] Koczkodaj, W. (2017). Somerville Happiness Survey [Dataset]. UCI Machine Learning Repository.

https://doi.org/10.24432/C5PW36.