Research on Hyperuricemia Diagnostic Model Based on Machine Learning and Infrared Thermography

Jun Wang, Danxuan Zhang, Chengzhi Liu and Xiaoling Zhou

***Abstract*—Hyperuricemia, a common metabolic disorder caused by impaired purine metabolism or reduced uric acid excretion, is a major cause of gout and a significant risk factor for cardiovascular disease and diabetes. This study proposes an efficient diagnostic model integrating infrared thermography (IRT) with machine learning to enable non-invasive, early detection. Using IRT, body surface temperatures from 17 anatomical regions were recorded in 262 hyperuricemic patients and 274 healthy controls. Data were normalized, analyzed for correlations, and reduced in dimensionality via principal component analysis (PCA). A random forest (RF) algorithm with optimized hyperparameters was then trained on the PCA features. The RF model demonstrated strong performance on both training and validation sets, confirming its robustness and reliability. This approach offers a novel, non-invasive method for early hyperuricemia diagnosis.**

***Index Terms*—hyperuricemia, diagnostic model, principal component analysis, random forest.**

I. Introduction

HYPERURICEMIA is a common metabolic disorder caused by impaired purine metabolism or reduced uric acid excretion, with its global incidence rising due to lifestyle and dietary changes. Beyond being a major contributor to gout, hyperuricemia is a significant risk factor for cardiovascular disease and diabetes, which can severely impair patients' quality of life and long-term health.

The condition is often insidious and progressive, with early symptoms that are subtle or absent. As a result, diagnosis frequently occurs only after the onset of overt symptoms, missing the optimal treatment window. Early detection is therefore crucial. Although traditional blood biochemical tests are accurate, they are invasive, time-consuming, and costly, limiting their suitability for timely or large-scale screening and consequently reducing the effectiveness of early intervention.

To overcome these limitations, researchers have investigated various strategies to improve diagnostic accuracy and efficiency. Beyond refinements in biochemical testing, emerging approaches include the use of novel biomarkers [1] and artificial intelligence (AI)-based tools capable of analyzing large-scale medical data to uncover patterns in hyperuricemia [2].

Non-invasive diagnostic methods avoid surgical or penetrating procedures, relying instead on external techniques to assess physiological states. They are increasingly valued for their safety, convenience, and improving diagnostic accuracy. Common non-invasive modalities include imaging, electrocardiography, endoscopy, functional magnetic resonance imaging, biomarker detection, and thermal imaging.

Several non-invasive approaches for hyperuricemia have been explored. For instance, Liang et al. developed prediction models using physical examination indicators and classification and regression trees [3]. Zhang et al. identified microbial biomarkers and demonstrated that combining clinical and microbial features yields a reliable diagnostic tool [1]. Shi et al. proposed a gender-specific model with distinct predictors for men and women, achieving satisfactory performance [4]. Shen et al. developed a model to improve early detection and predict gout progression, effectively distinguishing gout from asymptomatic hyperuricemia [5]. In hypertensive patients, Zhang et al. constructed a nomogram with strong discriminative ability for hyperuricemia, aiding cardiovascular risk assessment [1]. Lee et al. compared several machine learning algorithms with logistic regression, finding that Naive Bayes and RF performed best [6]. Chen et al. built a non-invasive prediction model for Chinese adults using modifiable risk factors, with XGBoost outperforming logistic regression and RF [7]. Gao et al. introduced an RF-based model to predict the risk of hyperuricemia [8]. In [9], Zeng et al. developed an artificial neural network based on dietary factors, demonstrating high predictive accuracy.

As a non-invasive modality, the infrared thermography (IRT) has wide-ranging applications in medical diagnosis. It offers rapid, painless, and non-contact imaging, detecting subtle temperature variations on the skin surface that may reflect underlying pathology. Machine learning has demonstrated substantial potential across diverse domains [10], [11], making it a promising complement to IRT for diagnostic purposes.

In this study, we analyze infrared temperature data with machine learning techniques, particularly RF, aiming to develop a highly accurate and efficient diagnostic model for hyperuricemia. To the best of our knowledge, this is the first work to combine IRT, PCA, and RF for the non-invasive diagnosis of hyperuricemia and to validate its

Manuscript received February 3, 2025; revised August 26, 2025.

This work was supported in part by the Guangxi University of Chinese Medicine Joint Fund Project under Grant 2023L2066.

J. Wang is an undergraduate student of School of Mathematics and Finance, Hunan University of Humanities, Science and Technology, Loudi 417000, P. R. China (e-mail: 1918838015@qq.com).

D. X. Zhang is an associate chief physician of Liuzhou Traditional Chinese Medical Hospital, Liuzhou 514001, P. R. China (e-mail: 279711128@qq.com).

C. Z. Liu is an associate professor of School of Mathematics and Finance, Hunan University of Humanities, Science and Technology, Loudi 417000, P. R. China (corresponding author to provide e-mail: it-rocket@163.com).

X. L. Zhou is a chief physician of Liuzhou Traditional Chinese Medical Hospital, Liuzhou 514001, P. R. China (corresponding author to provide e-mail: zxl-lz@163.com).

clinical feasibility.

II. Data acquisition

This study enrolled 536 participants, comprising 262 patients with hyperuricemia and 274 healthy controls. IRT was used to systematically collect body surface temperature data from predefined anatomical sites. This non-invasive method enabled the establishment of a comprehensive dataset for subsequent analysis.

The cohort had a relatively balanced gender distribution (319 males and 217 females), ensuring diversity and representativeness. Temperature measurements were obtained from 17 anatomical regions: both eyes, both axillae, both lateral thoracic regions, both palms, the dorsal surfaces of both feet, both renal areas, the Du meridian, and regions corresponding to traditional Chinese medicine meridians, including the upper jiao, heart, middle jiao, and lower jiao.

By analyzing these thermographic data, we aim to identify thermographic patterns and anomalies that aid in developing a diagnostic model for hyperuricemia.

III. Data preprocessing

A. Data normalization

To ensure comparability of body surface temperature measurements across different anatomical regions and maintain analytical consistency, a standardized preprocessing procedure was applied. This process transforms the data into a dimensionless form, mitigating the influence of scale differences and thereby enhancing the validity of inter-regional comparisons. The normalization formula is given by

$$X_{\text{norm}} = \frac{X - \mu}{\sigma},$$

where X represents the original dataset of body surface temperatures, μ is the mean of the dataset, and σ denotes the standard deviation.

B. Correlation analysis

Before performing principal component analysis (PCA [12], [13]), it is essential to conduct a correlation analysis of the normalized body surface temperature data to assess the suitability of PCA.

The normalized data were evaluated using the Kaiser-Meyer-Olkin measure (KMO [14], [15]) and Bartlett's test for sphericity [16], [17], with the results summarized in Table I. As shown in Table I, the KMO value of 0.932, which is close to the ideal value of 1, suggests strong partial correlations among variables, supporting the suitability of PCA. Additionally, Bartlett's test yielded a significance level of 0.000, well below the 0.05 threshold, further confirming sufficient correlations to justify PCA application.

C. Data dimension reduction

Next, PCA was applied to the normalized temperature data to reduce dimensionality, thereby simplifying data structure and extracting the most informative features.

TABLE I
Results of the KMO and Bartlett's tests.

KMO test	Bartlett's sphericity test		
	Approximate χ^2	Degrees of freedom	Significance
0.932	18550.738	136	0.000

In PCA, each principal component (PC) is associated with an eigenvalue representing the proportion of total variance explained. According to the Kaiser criterion ([18], [19]), principal components (PCs) with eigenvalues greater than 1 are considered significant, as they account for more variance than a single original variable. The eigenvalues of the first four PCs are 10.847, 1.702, 1.010, and 1.003 respectively. All of them exceed the threshold of 1. These components therefore collectively explain a substantial portion of the variance and were preserved for further analysis.

To further evaluate the cumulative explanatory power of the selected four PCs, we plotted their explained variance ratios, as shown in Fig. 1. We can see that the cumulative explained variance of the first four PCs exceeds 85%, indicating that these components effectively capture the majority of the data's variance and adequately represent the underlying structure.

The original normalized temperature data matrix X_{norm} (of size 536×17) was projected onto the PCA eigenvector matrix V (of size 17×4) to obtain the principal component scores matrix T (size 536×4), computed as

$$T = X_{\text{norm}} \times V,$$

where each column of T corresponds to the scores of a particular principal component for all subjects. These PC scores, reflecting distinct thermographic patterns, are presented in Table II and may possess clinical relevance.

These interpretable PC scores served as input features for training the random forest classification model. By preserving physiologically meaningful thermal patterns while effectively reducing dimensionality, the use of PC scores enhances both computational efficiency and diagnostic accuracy for hyperuricemia classification.

TABLE II
Eigenvector matrix of the first four PCs derived from PCA on body surface temperature data.

	PC 1	PC 2	PC 3	PC 4
Du Meridian	0.2947	-0.0047	0.0213	-0.0050
Left Eye	0.2978	-0.0011	-0.0048	-0.0175
Right Eye	0.2987	-0.0020	-0.0051	-0.0130
Left Axillary	0.2945	-0.0012	0.0194	-0.0068
Right Axillary	-0.0009	-0.0040	-0.5661	-0.7041
Upper Jiao	0.2959	-0.0061	0.0335	0.0006
Heart	0.2930	-0.0072	0.0321	-0.0109
Middle Jiao	0.2952	-0.0082	0.0430	0.0165
Lower Jiao	0.2970	-0.0116	0.0218	0.0094
Left Rib	0.2974	-0.0062	0.0249	0.0002
Right Rib	-0.0025	-0.0083	0.6432	-0.0746
Left Palm	0.2577	0.0175	-0.0851	-0.0014
Right Palm	0.0051	-0.0035	-0.4967	0.7053
Left Dorsal Foot	0.2709	0.0057	-0.0559	0.0027
Right Dorsal Foot	0.2673	0.0099	-0.0515	0.0138
Left Kidney	0.0013	0.7069	0.0037	-0.0008
Right Kidney	0.0065	0.7067	0.0033	-0.0006

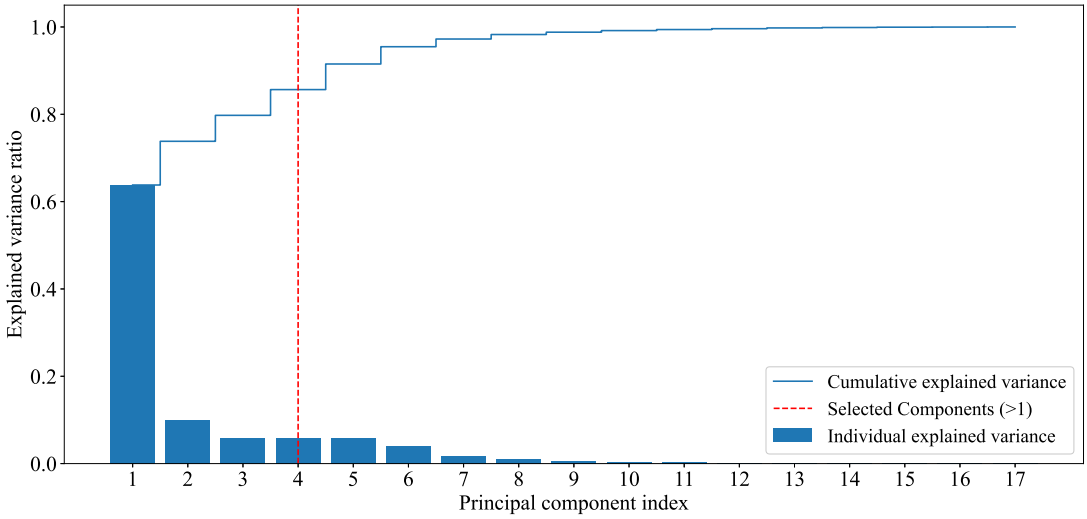

Fig. 1. Plot of the explained variance of the principal components.

IV. RF-BASED DIAGNOSTIC MODEL FOR HYPERURICEMIA

We developed the diagnostic model using the RF algorithm ([20], [21]) and compared its performance against XGBoost ([7]), LightGBM ([22]), SVM ([23], [24]), and KNN ([25], [26]). To ensure a rigorous and unbiased performance evaluation, the dataset was randomly partitioned into a training set and a validation set, with 70% of the data used for training and the remaining 30% for validation.

The RF-based diagnostic model for hyperuricemia was constructed through a systematic process, including bootstrap sampling, decision tree generation, and majority voting. As the performance of RF is highly sensitive to the choice of hyperparameters, a grid search strategy with cross-validation was employed to identify the optimal configuration. The hyperparameters of RF include: `max_depth` (the maximum depth of each decision tree, controlling model complexity), `max_features` (the number of features randomly selected at each split, enhancing generalization), `min_samples_leaf` (the minimum number of samples required at a leaf node, preventing overfitting to small subsets), `min_samples_split` (the minimum number of samples required to further split an internal node, ensuring sufficient partitioning), and `n_estimators` (the number of trees in the ensemble, improving stability and robustness).

The best-performing hyperparameters were then used to train the final RF diagnostic model. This optimization process enabled the model to strike a favorable balance between bias and variance, thereby reducing the risk of overfitting. The complete modeling pipeline is summarized in Algorithm 1.

Fig. 2 illustrates the overall flowchart of the machine learning pipeline for hyperuricemia diagnosis. The pipeline comprises four main stages: (i) data preprocessing, where data normalization and PCA-based dimensionality reduction are performed; (ii) data splitting, where the dataset is partitioned into training and validation subsets; (iii) model training, during which classifiers such as RF, SVM, KNN, XGBoost, and LightGBM are trained and compared; and (iv) prediction, in which the best-performing model is used to generate the final diagnosis together with associated probability scores.

In our experiments, a grid search with 5-fold cross-validation was performed on the training dataset, using accuracy and the area under the ROC curve (AUC) as evaluation metrics. The optimal hyperparameter configuration was determined as follows:

`max_depth=20`,
`max_features=2`,
`min_samples_leaf=1`,
`min_samples_split=12`,
`n_estimators=500`.

To evaluate the performance of our model, we present the receiver operating characteristic (ROC) curves and the AUC values for the RF, SVM, KNN, XGBoost, and LightGBM algorithms in Fig. 3. The ROC curve is a crucial tool for assessing classification performance, as it illustrates the trade-off between the true positive rate and the false positive rate, thereby reflecting the model's discriminative capability. The dashed line in the graph indicates the baseline for random guessing; the closer the curve is to the upper left corner, the better the model's performance. The AUC provides a quantitative measure of overall classification effectiveness, with values closer to 1 indicating superior model performance. The AUC values for the RF, SVM, KNN, XGBoost, and LightGBM algorithms are 0.75, 0.70, 0.66, 0.68, and 0.65, respectively, indicating that although performance differences are modest, the RF algorithm achieves the highest discriminative ability among the three methods.

Furthermore, Fig. 4 compares the precision, recall, and F1-score, where healthy individuals are labeled as 0 and hyperuricemic patients as 1. Except for a slightly higher recall for label 0 by KNN, RF consistently outperforms SVM, KNN, XGBoost, and LightGBM, showing particularly strong performance in correctly identifying hyperuricemic patients. These results demonstrate that RF excels in AUC, precision, recall, and F1-score, confirming its efficiency and robustness as a diagnostic tool for hyperuricemia detection.

Algorithm 1 RF-Based Diagnostic Model for Hyperuricemia.

1: **Hyperparameter optimization (grid search)**: Perform the grid search strategy on the training set to identify the optimal hyperparameters that maximize accuracy and AUC. The best configuration is recorded and used for subsequent model construction.
2: **Bootstrap sampling**: Perform `n_estimators` rounds of bootstrap sampling on the training set to create `n_estimators` different subsets. Each subset is used to build an individual decision tree, thereby improving model diversity and generalization.
3: **Feature selection**: For each decision tree, randomly select `max_features` features (out of the total M features) for consideration at each node split.
4: **Node splitting**: Recursively split each node using the optimal feature and threshold that maximize data purity until stopping criteria are met (controlled by `max_depth`, `min_samples_split`, and `min_samples_leaf`).
5: **Tree construction**: Build a collection of diverse and independent decision trees to capture different aspects of the data.
6: **RF model establishment**: Combine the decision trees into an ensemble model using majority voting. The final classification result for a sample D is given by

$$R(D) = \begin{cases} 0, & \text{No hyperuricemia;} \\ 1, & \text{With hyperuricemia.} \end{cases}$$

7: **Performance evaluation**: Evaluate the model using standard classification metrics. Let TP, FP, and FN denote the number of true positives, false positives, and false negatives, respectively. Then

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}},$$
$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}},$$
$$\text{F1} = 2\frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}.$$

8: **Model validation**: Validate the model performance on the held-out validation set by comparing predicted outcomes with actual labels.
9: **Result analysis and model optimization**: If performance metrics meet the desired threshold, the model is finalized. Otherwise, adjust parameters or try different methods, and the training-validation process is repeated until satisfactory performance is achieved.

V. Conclusions

In this paper, we developed a diagnostic model for hyperuricemia by integrating infrared thermography with machine learning techniques, particularly the RF algorithm. The proposed model achieved robust and consistent performance in both the training and validation phases, effectively discriminating between hyperuricemic patients and healthy individuals. Its non-invasive non-invasive, rapid, and cost-effective characteristics highlight its potential as a practical tool for early clinical diagnosis, thereby supporting improved patient management and outcomes.

Future research will aim to validate the model's generalizability across larger and more diverse populations, and to explore advanced feature selection and algorithmic optimization strategies to further enhance diagnostic accuracy.

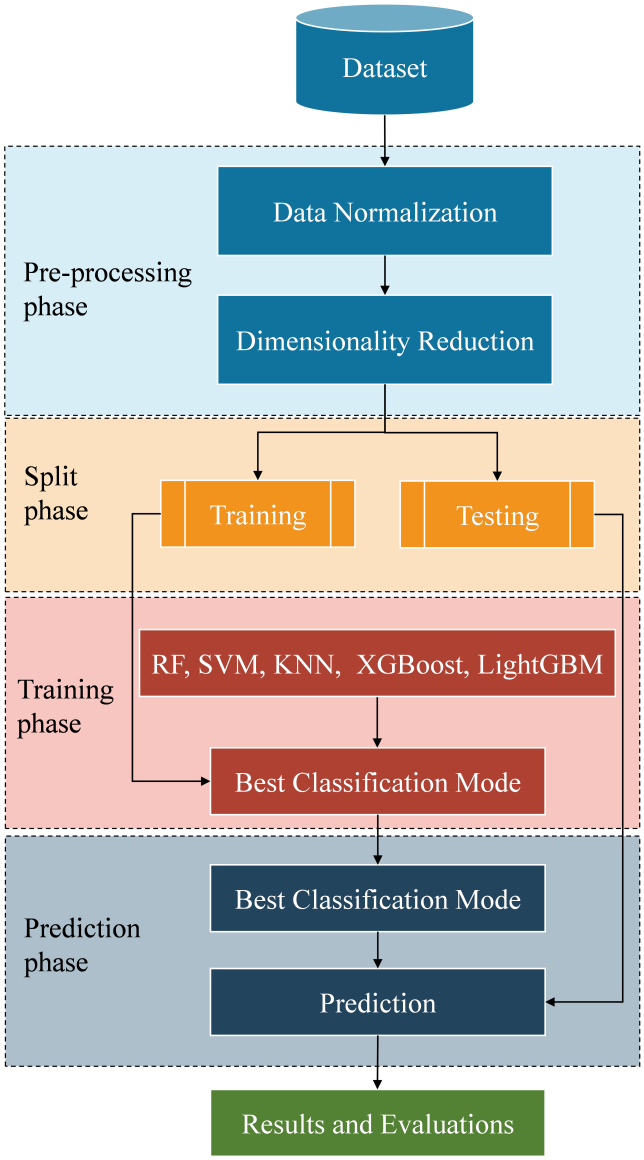

Fig. 2. Flowchart of machine learning-based diagnostic models for hyperuricemia.

References

[1] L. X. Zhang, J. Y. Cao and X. J. Zhou, "Development and validation of a prediction model for hyperuricemia risk in hypertensive patients," *American Journal of Cardiovascular Disease*, vol. 14, no. 1, pp1-8, 2024.
[2] Z. X. Chen, W. Wang, H. Tian, W. R. Yu, Y. Niu, X. L. Zheng, S. H. Liu, L. Wang and Y. Z. Huang, "Wearable intelligent sweat platform for SERS-AI diagnosis of gout," *Lab on a Chip*, vol. 24, no. 7, pp1996-2004, 2024.
[3] M. T. Liang, J. K. Liu, W. J. Chen, Y. He, M. Kahaer, R. Li, T. T. Tian, Y. Z. Liu, B. Bai, Y. N. Cui et al., "Diagnostic model for predicting hyperuricemia based on alterations of the gut microbiome in individuals with different serum uric acid levels," *Frontiers in Endocrinology*, vol. 13, pp925119, 2022.

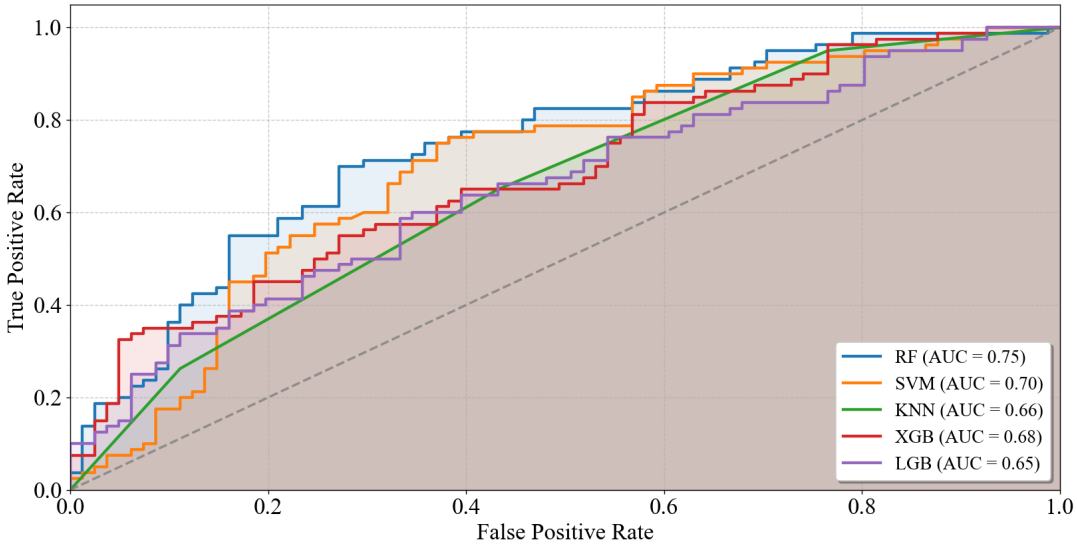

Fig. 3. Comparison of ROC curves for the RF, SVM, KNN, XGBoost, and LightGBM algorithms.

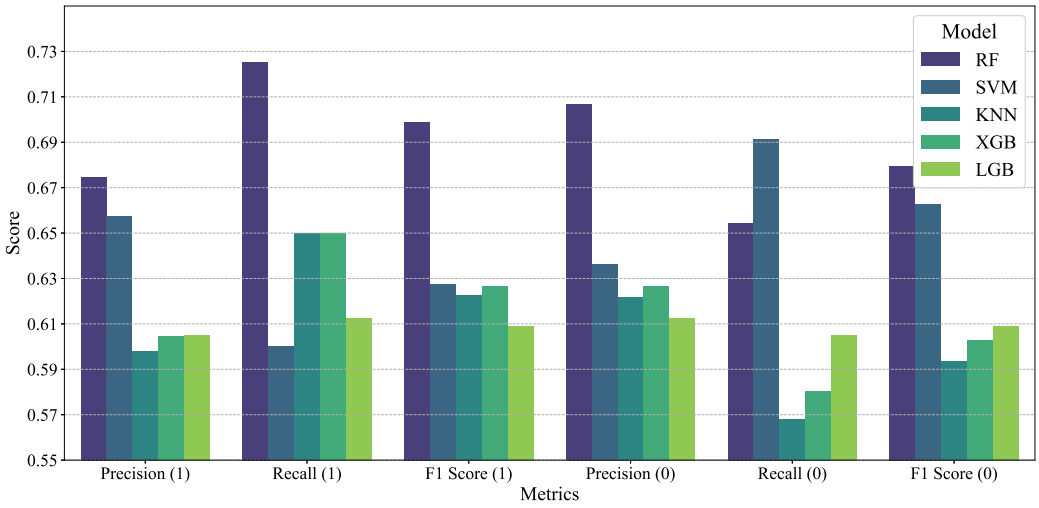

Fig. 4. Comparison of the precision, recall, and F1 score for the RF, SVM, KNN, XGBoost, and LightGBM algorithms.

[4] J. C. Shi, X. H. Chen, Q. Yang, C. M. Wang, Q. Huang, Y. M. Shen and J. Yu, "A simple prediction model of hyperuricemia for use in a rural setting," *Scientific Reports*, vol. 11, no. 1, pp23300, 2021.

[5] X. Shen, C. Wang, N. N. Liang, Z. Liu, X. D. Li, Z. J. Zhu, T. R. Merriman, N. Dalbeth, R. Terkeltaub and C. G. Li et al., "Serum metabolomics identifies dysregulated pathways and potential metabolic biomarkers for hyperuricemia and gout," *Arthritis & Rheumatology*, vol. 73, no. 9, pp1738-1748, 2021.

[6] S. Lee, E. K. Choe and B. Park, "Exploration of machine learning for hyperuricemia prediction models based on basic health checkup tests," *Journal of Clinical Medicine*, vol. 8, no. 2, pp172, 2019.

[7] S. Chen, W. Han, L. R. Kong, Q. Li, C. D. Yu, J. B. Zhang and H. J. He, "The development and validation of a non-invasive prediction model of hyperuricemia based on modifiable risk factors: baseline findings of a health examination population cohort," *Food & Function*, vol. 14, no. 13, pp6073-6082, 2023.

[8] Y. H. Gao, S. C. Jia, D. H. Li, C. Huang, Z. W. Meng, Y. Wang, M. Yu, T. Y. Xu, M. Liu and J. H. Sun et al., "Prediction model of random forest for the risk of hyperuricemia in a Chinese basic health checkup test," *Bioscience Reports*, vol. 41, no. 4, ppBSR20203859, 2021.

[9] J. Zeng, J. G. Zhang, Z. Y. Li, T. W. Li and G. W. Li, "Prediction model of artificial neural network for the risk of hyperuricemia incorporating dietary risk factors in a Chinese adult study," *Food & Nutrition Research*, vol. 64, no. 3712, 2020.

[10] M. Abed, M. A. Imteaz and A. N. Ahmed, "Utilising machine learning for pan evaporation prediction - A case study in western australia," *Lecture Notes in Engineering and Computer Science: Proceedings of the International MultiConference of Engineers and Computer Scientists*, Hong Kong, pp14-18, 2023.

[11] A. Ayub Syed, Y. Heryadi, Lukas, and A. Wibowo, "A comparison of machine learning classifiers on laptop products classification task," *Lecture Notes in Engineering and Computer Science: Proceedings of the International MultiConference of Engineers and Computer Scientists 2021*, Hong Kong, pp104-110, 2021.

[12] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 4, pp433-459, 2010.

[13] I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 374, no. 2065, pp20150202, 2016.

[14] H. F. Kaiser, "A second generation little jiffy," *Psychometrika*, vol. 35, no. 4, pp401-415, 1970.

[15] H. F. Kaiser and J. Rice, "Little jiffy, mark IV," *Educational and Psychological Measurement*, vol. 34, no. 1, pp111-117, 1974.

[16] M. S. Bartlett, "Properties of sufficiency and statistical tests," *Proceedings of the Royal Society of London. Series A-Mathematical and Physical Sciences*, vol. 160, no. 901, pp268-282, 1937.

[17] H. Arsham and M. Lovric, "Bartlett's Test," *International Encyclopedia of Statistical Science*, vol. 2, pp20-23, 2011.

[18] R. Larsen and R. T. Warne, "Estimating confidence intervals for eigenvalues in exploratory factor analysis," *Behavior Research Methods*, vol. 42, pp871-876, 2010.

[19] R. T. Warne and R. Larsen, "Evaluating a proposed modification of the Guttman rule for determining the number of factors in an exploratory

factor analysis," *Psychological Test and Assessment Modeling*, vol. 56, no. 1, pp104-123, 2014.
[20] L. Breiman, "Random forests," *Machine Learning*, vol. 45, pp5-32, 2001.
[21] S. J. Rigatti, "Random forest," *Journal of Insurance Medicine*, vol. 47, no. 1, pp31-39, 2017.
[22] D. Meng, H. Dai and Q. Sun et al, "Novel Wireless Sensor Network Intrusion Detection Method Based on LightGBM Model," *IAENG International Journal of Applied Mathematics*, vol. 52, no.4, pp955-961, 2022.
[23] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt and B. Scholkopf, "Support vector machines," *IEEE Intelligent Systems and Their Applications*, vol. 13, no. 4, pp18-28, 1998.
[24] D. A. Pisner and D. M. Schnyer, "Support vector machine," *Machine Learning*, pp101-121, 2020.
[25] J. Laaksonen and E. Oja, "Classification with learning k-nearest neighbors," *Proceedings of International Conference on Neural Networks*, vol. 3, pp1480-1483, 1996.
[26] L. E. Peterson, "K-nearest neighbor," *Scholarpedia*, vol. 4, no. 2, pp1883, 2009.