BMC Medical Informatics and Decision Making

RESEARCH

Open Access

Can some algorithms of machine learning identify osteoporosis patients after training and testing some clinical information about patients?

Guixiong Huang[1*†], Weilin Zhu[2†], Yulong Wang[3], Yizhou Wan[1], Kaifang Chen[1], Yanlin Su[1], Weijie Su[1], Lianxin Li[4], Pengran Liu[1*] and Xiao dong Guo[1*]

Abstract

Objective This study was designed to establish a diagnostic model for osteoporosis by collecting clinical information from patients with and without osteoporosis. Various machine learning algorithms were employed for training and testing the model, evaluating its performance, and conducting validations to explore the most suitable machine learning algorithm.

Methods Clinical information, including demographic data, examination results, medical history, and laboratory test results, was collected from inpatients with and without osteoporosis. The LASSO algorithm was utilized for feature selection, and multiple machine learning algorithms were applied to calculate the model's accuracy, precision, recall, F1 score, and average precision (AP) value. Receiver operating characteristic (ROC) curves for each algorithm were plotted, and a comprehensive evaluation was conducted to identify the most suitable machine learning model. Finally, the model's predictive accuracy was validated using corresponding information from other patients.

Results A total of 1063 patients were included; 562 had osteoporosis, and 501 did not. After LASSO feature selection, the most important features for the model's predictive results were determined to be age, height, weight, alkaline phosphatase activity, and osteocalcin. Evaluation of the accuracy, precision, recall, F1 score, and AP value for each algorithm, along with ROC curves, led to the selection of the light gradient boosting machine (LGBM) algorithm as the best algorithm for the model. The validation results confirmed the model's excellent predictive ability.

†Guixiong Huang and Weilin Zhu have contributed equally in the planning and writing of the manuscript as the first author.

Xiao dong Guo is the main corresponding author of this article.

*Correspondence:
Guixiong Huang
huangguixiong@hust.edu.cn
Pengran Liu
lprlprlprwd@163.com
Xiao dong Guo
xiaodongguo@hust.edu.cn

Full list of author information is available at the end of the article

Conclusion This study established a preliminary diagnostic model for osteoporosis, contributing to increased efficiency in diagnosing the disease.

Keywords Osteoporosis, Machine learning, Fragility fracture

Background

Osteoporosis is a common and potentially slow-developing disease associated with degeneration of the body [1]. For hospitalized patients, there may be no early symptoms [2]. Therefore, clinical doctors may easily overlook the diagnosis of such patients. When osteoporosis progresses to a certain extent, patients may experience fragility fractures in the spine, hips, or wrists with minimal or no apparent external force [3]. Hence, treatment for osteoporosis and fragile bone diseases is essential. However, the treatment of osteoporosis and the occurrence of fragility fractures impose substantial economic burdens on patients' families and society [4]. To reduce economic costs, early prediction and prevention of osteoporosis are crucial. This study focused on identifying a method for assessing the risk of osteoporosis.

The diagnosis of osteoporosis can be overlooked by nonspecialized medical professionals. To address this issue, we collected partial clinical data and data from hospitalized patients as training data. By utilizing machine learning methods [5, 6], the data were trained and validated with the aim of maximizing the probability of diagnosing osteoporosis in hospitalized patients. This study adopts a novel approach compared to conventional methods for diagnosing osteoporosis. The use of machine learning is designed to maximize the identification of individuals with this disease, providing diagnostic criteria for regions with limited access to osteoporosis detection instruments, remote areas, and doctors unfamiliar with the disease.

Method

This study collected data from the Tongji Medical College Affiliated Union Hospital of Huazhong University of Science and Technology and Shenzhen Health Development Research and Data Management Center over the time period of October 2019 to December 2021. The collected data included information on the presence of osteoporosis, age, sex, height, weight, body mass index (BMI), smoking status, hypertension status, diabetes status, history of trauma, history of osteoporosis, alkaline phosphatase activity, blood calcium, blood phosphorus, urine protein, and osteocalcin.

The inclusion criteria for patients were as follows: age ≥ 30 years, primary osteoporosis, and a control group excluding patients with osteoporosis.

The exclusion criteria for patients were as follows: age < 30 years, had a tumour, was malnourished, had an autoimmune disease, was a long-term oral steroid user, was receiving chemotherapy, or had other conditions affecting bone density.

The collected information was assigned numerical values for representation. Osteoporosis patients were assigned a value of 1; otherwise, 0; males were assigned 1; females, 0; height was measured in centimetres (cm); weight was measured in kilograms (kg); smoking was assigned 1; otherwise, 0; presence of hypertension, diabetes, history of osteoporosis, and history of trauma were assigned 1; otherwise, 0; alkaline phosphatase activity was measured in U/L, assigned 0 if not checked; blood calcium concentration was measured in mmol/L, assigned 0 if not checked; blood phosphorus concentration was measured in mmol/L, assigned 0 if not checked; osteocalcin was measured in ng/ml, assigned 0 if not checked; and urine protein positivity was assigned 1 if negative. These data were then organized into an Excel spreadsheet.

Some of the information collected did not contribute significantly to predicting whether a patient had osteoporosis. To simplify the model, the least absolute shrinkage and selection operator (LASSO) algorithm [7] was used for feature selection. The mean squared error (MSE) was employed in this study to handle the model's feature coefficients, finding the optimal regularization parameter λ (lambda) corresponding to the MSE. This helped the model achieve appropriate complexity and generalization performance, and the process was visualized. The relationship between the optimal lambda and the corresponding model coefficients was also visualized. Since the collected data could inconsistent units, leading to large differences in magnitude across variables, data preprocessing was conducted via normalization (also known as standardization) to make the data comparable, following the methods of several scholars [8]. In this study, the dataset was split into a training set and a test set with a ratio of 0.7 to 0.3, respectively. Additionally, five-fold cross-validation was performed to ensure the robustness of the model.

In accordance with the methods of many scholars, the computational methods used included support vector machines (SVMs) [9–12], stochastic gradient descent (SGD) [10], K-nearest neighbour (KNN) methods [11], decision tree (DT) methods [12], random forest (RF) methods [11, 12], extremely randomized trees (ERs) [13], eXtreme Gradient Boosting (XGB) methods [14], light gradient boosting machines (LGBMs) [15], and logistic regression (LR) methods [10, 12, 16]. The accuracy, precision, recall, F1 score, AP value, and ROC curve [17] of the

Table 1 Patient information for 9 patients

	Patient's number								
Categories	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Diagnosis	Y	Y	N	Y	N	Y	N	N	Y
Sex	F	F	F	F	F	F	M	M	F
Age, year	76	74	66	70	51	77	55	58	88
Height, cm	160	155	158	159	160	160	174	175	155
Weight, kg	65	55	56.5	60	70	55	85	80	55
Body mass index	23.44	22.89	22.63	23.73	27.34	21.48	28.08	26.12	22.89
Smoking	N	N	N	N	N	N	N	N	N
Hypertension	N	N	Y	N	Y	Y	N	N	Y
diabetes	N	N	Y	N	N	N	Y	N	N
Trauma	N	N	N	Y	N	Y	N	Y	Y
History of osteoporosis	N	N	N	Y	N	Y	N	Y	Y
Alkaline phosphatase, U/L	65	193	46	71	50	58	60	99	72
Calcium, mmol/L	2.2	2.21	2.06	2.08	2.19	3.24	2.15	2.21	2.25
Phosphorus, mmol/L	0.91	0.94	0.96	0.93	0.92	1.16	0.72	1.12	0.82
Proteinuria	-	-	-	-	-	-	-	-	+
Osteocalcin, ng/ml	9.49	22.77	ND	ND	ND	27.06	ND	ND	10.14

Note: N, no; Y, yes; -, negative; +, positive; ND, not done

aforementioned models were calculated. Comprehensive evaluation and weighting were performed to determine the best computational method.

We used data from 5 patients with osteoporosis and 4 patients without osteoporosis (as shown in Table 1) and validated them using the best computational method. The aim was to observe whether this computational method could accurately distinguish whether these patients had osteoporosis. To provide insight into the model's predictions, this study applied Local Interpretable Model-agnostic Explanations (LIME) [18] to analyze the contributing factors.

Statistical methods

This study employed t tests for statistical comparisons, where $P<0.05$ indicated statistical significance. All algorithms and statistical methods were implemented using Python 3.8 on the Jupyter Notebook platform for analysis and computation.

Results

A total of 1063 patient records were collected in this study, including 562 patients with osteoporosis and 501 randomly selected patients without osteoporosis. Among them, 745 were female and 318 were male, with an age range of 33 to 94 years and an average age of 62.62 ± 12.78 years. The height ranged from 140.0 to 183.0 cm, with an average height of 161.23 ± 6.99 cm. The weight ranged from 31.9 to 106 kg, with an average weight of 59.83 ± 11.06 kg. The BMI ranged from 11 to 27, with an average of 22.95 ± 3.56. Proteinuria was negative in 1008 patients and positive in 55 patients. Regarding history of osteoporosis, 209 had a history of osteoporosis, while 854 had no history of osteoporosis.

Table 2 Demographic characteristics of the participants

Variable	Osteoporosis	Nonosteoporosis	*P* value
Gender, No. (%)	562(52.87)	501(47.13)	0.01 < P
Age, mean (SD), year	62.62(12.78)	57.04(12.12)	0.01 < P
Height, mean (SD), mm	159(5.86)	163.50(7.26)	0.01 < P
Alkaline phosphatase, mean (SD), U/L	69(32.56)	72.18(30.54)	$P=0.21$
Weight, mean (SD), kg	57(8.96)	63.59(11.76)	0.01 < P
BMI, mean (SD)	22(3.23)	23.72(3.68)	0.01 < P
Proteinuria	24(4.27)	31(6.19)	-
Smoking, No. (%)	22(3.91)	70(13.97)	-
Hypertension, No. (%)	176(31.32)	180(35.93)	-
Diabetes, No. (%)	100(17.79)	88(17.65)	-
Trauma, No. (%)	103(18.33)	47(9.38)	-
History of osteoporosis, No. (%)	209(37.19)	0(0.00)	-

Note: SD, standard deviation

Most patients denied a history of smoking, hypertension, trauma, or past osteoporosis, as shown in Table 2. Regarding calcium, 22 patients did not undergo the test and were assigned a value of 0; regarding phosphorus, 471 patients did not undergo the test and were assigned a value of 0. Osteocalcin was not tested in 971 patients, and a value of 0 was assigned in those cases; true values were recorded for 55 patients.

After LASSO feature selection, the remaining features that contributed significantly to the predictive model

were age, height, weight, alkaline phosphatase activity, and osteocalcin. These features were included in the next steps of model training and validation. During feature selection, the regularization parameter lambda was adjusted through the mean squared error (MSE) process, and the optimal coefficient was found to be 0.08 as shown in Fig. 1. The relationship between the feature coefficients of the model and lambda changes is illustrated in Fig. 2.

The accuracy, precision, recall, F1 score, and AP value for each algorithm are presented in Table 3. According to Table 3, the accuracy ranges from 0.68 to 0.78, the precision ranges from 0.65 to 0.80, the recall ranges from 0.69 to 0.93, the F1 score ranges from 0.70 to 0.80, and the AP value ranges from 0.64 to 0.75.

In this study, the ROC curves for various computational methods are consistently close to the upper-left corner. The area under the curve (AUC) percentages are as follows: 0.83 for SVM, 0.79 for SGD, 0.82 for KNN, 0.73 for DT, 0.84 for RF, 0.86 for ET, 0.85 for XGB, 0.84 for LGBM, and 0.81 for LR. The ROC curves for each algorithm are illustrated in Fig. 3.

After comprehensively evaluating the accuracy, precision, recall, F1 score, and AP value for each algorithm, we combined the values into a composite score with weighted sums. The results were as follows: SVM, 3.85; SGD, 3.66; KNN, 3.75; DT, 3.65; RF, 3.93; ET, 3.85; XGB, 3.85; LGBM, 3.94; and LR, 3.78. The algorithm with the highest score was selected: LGBM. Subsequently, using the features of patients from Table 1 (features selected by Lasso in the model) to validate the predictive ability of the LGBM algorithm, the results showed successful predictions for 8 out of 9 patients. The 4th patient was misdiagnosed, a patient without osteoporosis being incorrectly classified as having osteoporosis as shown in Fig. 4.

The prediction results of the model were partially explained using LIME, as shown in Fig. 5 and Supplementary Material 1. As illustrated, the factor 'osteocalcin' had a significant contribution to the predictive model.

Discussion

In this study, the machine learning algorithms used demonstrated high accuracy, all exceeding 0.65. To enhance the predictive accuracy of the model, one can consider increasing the collected patient information or the number of cases. Some studies may employ data augmentation to increase the dataset size, thereby improving accuracy to a certain extent [19]. When using machine learning as a research method, many researchers [20, 21] commonly use metrics such as accuracy, precision, recall, F1 score, and AP value to assess the model's performance, where higher values indicate a better ability to correctly predict categories.

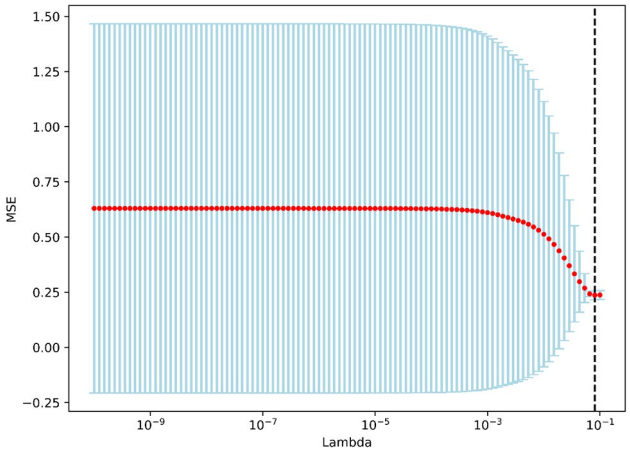

Fig. 1 Relationship between MSE and lambda (the blue bars indicate the range of standard deviation, and the black dashed line represents the optimal lambda value of 0.08, marked by the lowest position of the red dot)

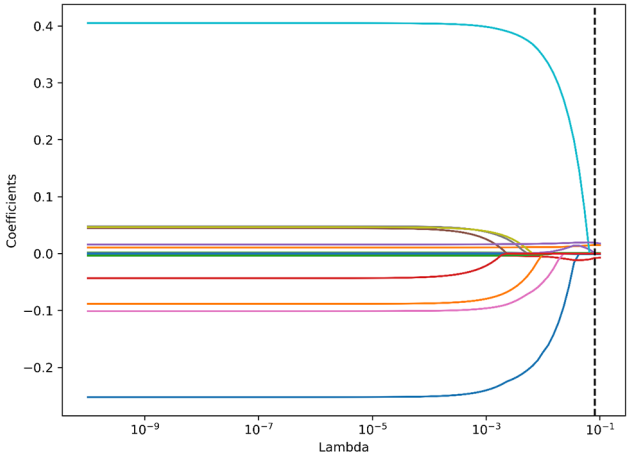

Fig. 2 Relationship between model coefficients and lambda (As shown in the figure, as lambda increases, the coefficients of certain features gradually become 0, while some feature coefficients remain unchanged; features with nonzero coefficients contributed to the diagnosis of osteoporosis in this study; this allows for the selection of features that are effective for prediction; the optimal lambda value in the figure is indicated by the dashed line at the intersection of the feature curve, which is 0.08.)

Table 3 Accuracy, precision, recall, F1 score, and AP value for each algorithm

Algorithms	Accuracy	Precision	Recall	F1 Value	AP Value
SVM	0.76	0.76	0.82	0.79	0.72
SGD	0.68	0.65	0.93	0.76	0.64
KNN	0.74	0.74	0.80	0.77	0.70
DT	0.72	0.78	0.69	0.73	0.71
RF	0.78	0.80	0.80	0.80	0.75
ET	0.76	0.79	0.78	0.78	0.74
XGB	0.76	0.78	0.79	0.79	0.73
LGBM	0.78	0.79	0.82	0.80	0.75
LR	0.75	0.78	0.86	0.77	0.72

Note: AP Value, Average Precision Value

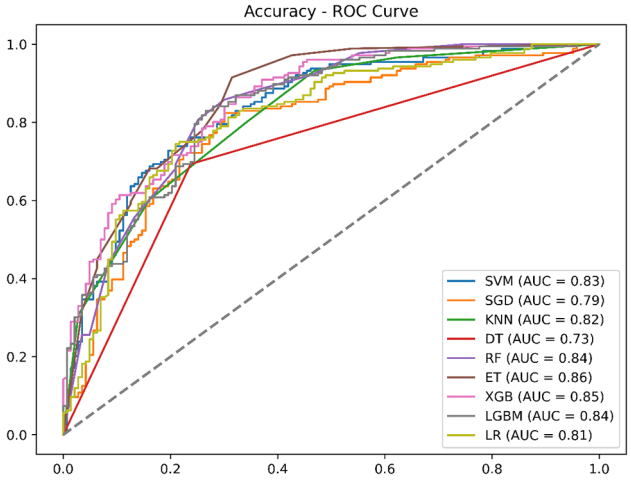

Fig. 3 ROC curves and corresponding AUC percentages for each algorithm (The dashed line represents the baseline of the ROC curve, signifying that the closer the model's performance is to this line, the weaker its predictive ability.)

In this study, age was found to play a significant role in predicting osteoporosis incidence. These findings confirm that the risk for postmenopausal osteoporosis increases in elderly women after menopause [22]. Therefore, for postmenopausal elderly female patients complaining of symptoms resembling osteoporosis, vigilance should be given regarding the possible presence of osteoporosis, and timely and relevant examinations should be performed for confirmation.

The retention of alkaline phosphatase activity after feature selection indicated its significant role in predicting whether patients had osteoporosis in this study. Alkaline phosphatase (ALP) is a bone formation marker [23] commonly used as an indicator of osteoporosis. Moreover, the level of this marker tends to increase in patients with osteoporosis.

This study suggested that height and weight are also factors that may be used to predict the risk of developing osteoporosis. These two factors can help minimize the risk of fragility fractures caused by osteoporosis [24]. Osteocalcin, the most abundant bone protein in bone cells, is used as a biochemical marker for bone [25]. In this study, features were retained after feature selection using the LASSO algorithm. Although some patients did not undergo osteocalcin testing, the LASSO algorithm in this study still effectively identified this factor, further validating the clinical relevance of the study results. LASSO retains the original features that are most predictive of the outcome, maintaining the clinical relevance and interpretability of the selected predictors. This characteristic is particularly critical in fields like healthcare, where understanding the role of individual features (e.g., BMI, smoking status) is as important as achieving high predictive accuracy.

Sex plays a role in predicting osteoporosis incidence. While the majority of patients in this study were female, the weight of sex in the prediction was not high after feature selection. The study suggested that although postmenopausal women often experience osteoporosis, not every postmenopausal woman will develop osteoporosis. Since data were collected from patients aged 30 years and older, the inclusion criteria may influence the model's ability to predict sex, leading to potential biases.

Several studies have indicated that BMI, smoking status, diabetes status, serum calcium concentration, serum phosphorus concentration, and proteinuria are associated with osteoporosis [25–28]. In this study, fewer patients with a history of smoking, diabetes, or proteinuria were recorded, and these factors were removed after

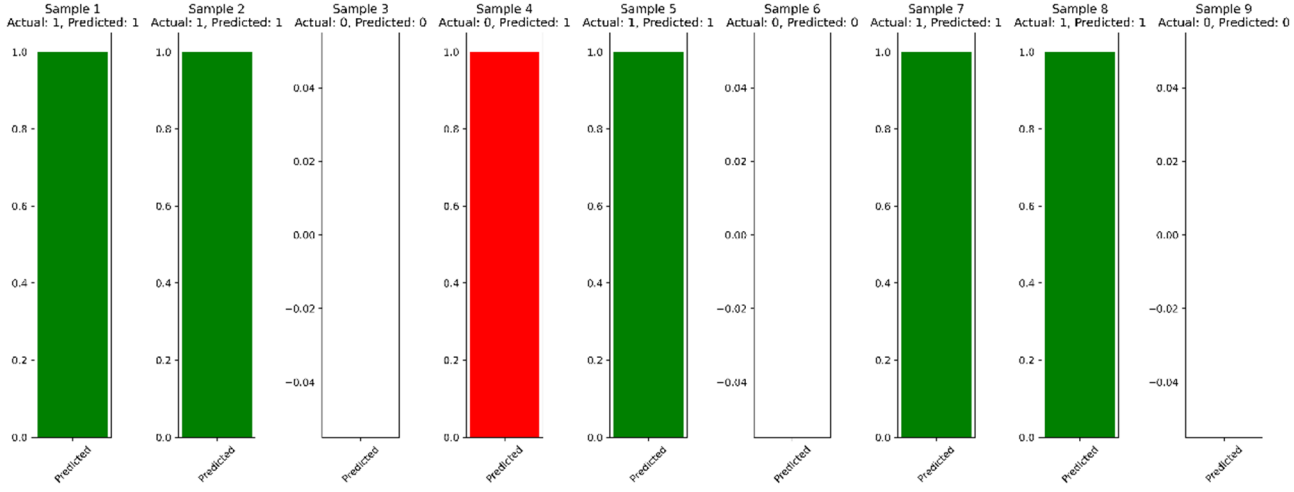

Fig. 4 Validation results of the LGBM algorithm for 9 patients (as shown in the figure, red indicates prediction errors, white and green indicate correct predictions, 0 represents patients without osteoporosis, and 1 represents patients with osteoporosis)

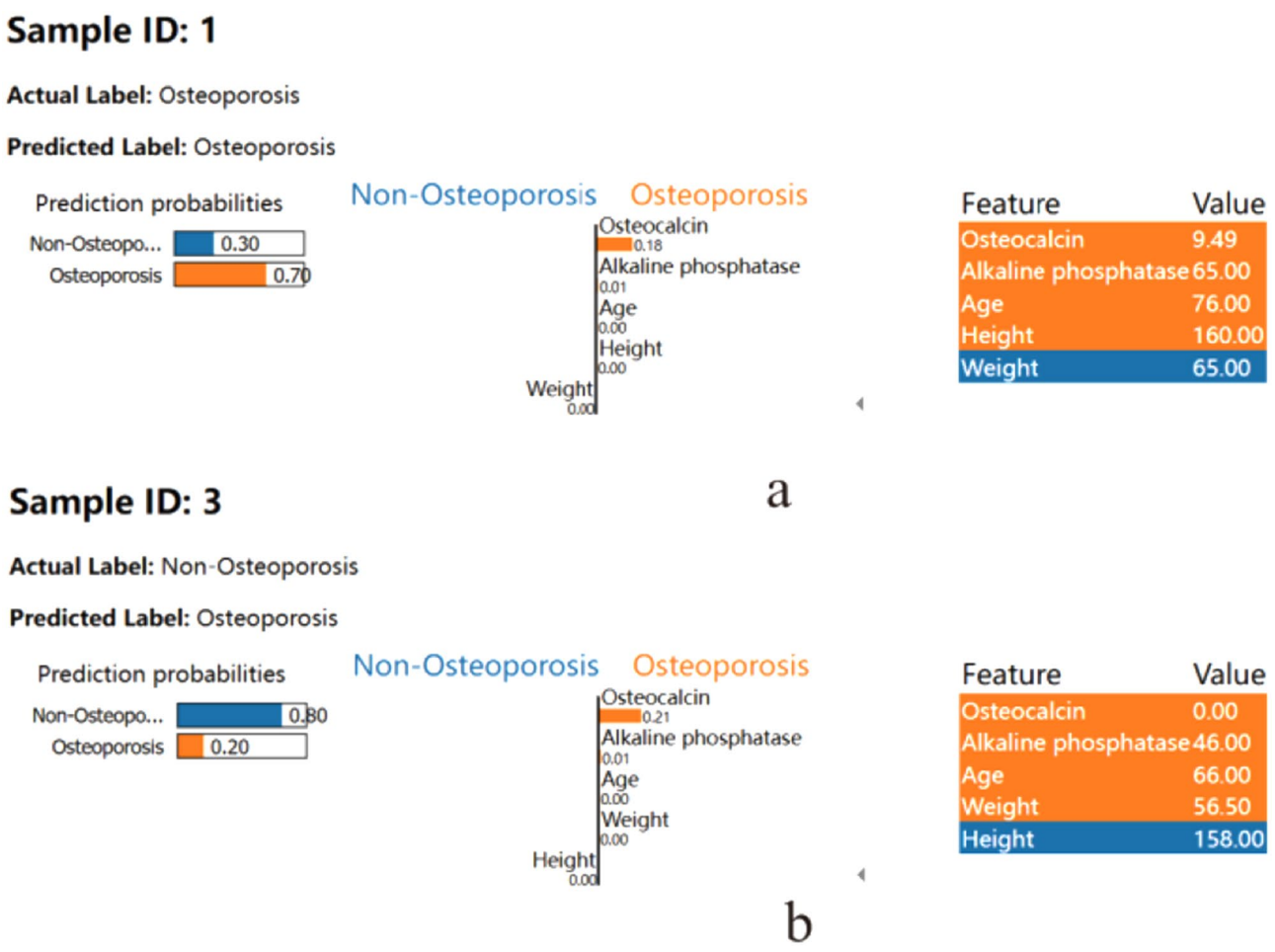

Fig. 5 Prediction explanations using LIME for Sample 1 and Sample 3 are presented. (**a**) For Sample 1, the prediction aligns accurately with the true label; (**b**) For Sample 3, the prediction does not match the true label, indicating an incorrect classification

feature selection, suggesting a relatively weak association with osteoporosis. This may be attributed to the data distribution in this study. In other words, in this study, the probability of these factors was roughly the same in patients with and without osteoporosis, resulting in LASSO failing to capture their impact on osteoporosis.

The ROC curves of the various algorithms in this study are closer to the upper-left corner, indicating larger areas under the curve and stronger predictive capabilities. These findings suggested that the selected algorithms in this study could be used to distinguish patients with osteoporosis from those without osteoporosis. After a comprehensive evaluation of the accuracy, precision, recall, F1 score, and AP score for each algorithm, the best-performing algorithm was identified as LGBM. In addition to achieving high predictive performance, LGBM inherently evaluates feature importance, which aligns with our goal of identifying the most relevant predictors in the dataset. This interpretability further justifies its use in this study, where understanding feature contributions is critical.

Validation with partial information from 9 patients showed that the LGBM algorithm could accurately predict patient type (accuracy of 88.89%). This further confirmed that the LGBM algorithm is suitable for classification recognition after training on the study data. This study recommends the use of this algorithm for predicting diseases in clinical research.

Although the SVM, SGD, KNN, DT, RF, ET, XGB, and LR algorithms can predict osteoporosis patients to some extent, they are not the optimal prediction algorithms according to comprehensive scores. This could be attributed to the relatively limited data in this study. To enhance predictive capabilities, increasing the number of patients, ensuring comprehensive information for each patient, and spanning a broader timeframe during patient enrolment are recommended. Collaborative research involving multiple centres could further maximize the accuracy of the research results.

Improving accuracy could benefit primary hospitals and remote clinics by enhancing their ability to diagnose osteoporosis and reducing potential risks associated

with fragility fractures. Establishing a reliable method for diagnosing osteoporosis in such medical institutions is crucial and was the primary goal of this study.

Therefore, we believe that with a sufficiently large database, a more accurate osteoporosis prediction model could be developed. This model could be integrated into hospital electronic medical record (EMR) systems as a plug-in or program. The trained model would be saved in a compatible format and packaged as an application programming interface (API) for easy integration. After compatibility testing with the hospital's server, the model would be embedded into the EMR system. The system would then be deployed with the integrated model, ensuring proper functionality, and periodic updates would be scheduled to maintain performance. By automatically identifying key predictive factors and assessing whether a patient has osteoporosis or the likelihood of developing it, this system could aid clinicians in making timely diagnoses and treatment decisions.

The prediction accuracy of the model in this study is high (8/9), although there are some inaccuracies (1/9). As shown in Fig. 5, sample 3 was misclassified. According to the feature contributions provided by LIME, osteocalcin had a relatively significant impact on the predicted outcome. In this sample, the factors osteocalcin, alkaline phosphatase, age, and weight (orange bars) all support the diagnosis of osteoporosis and are identified as risk factors. In contrast, the factor height (blue bar) opposes the diagnosis, acting as a protective factor. It is likely that the combined effect of the risk factors outweighs the protective factor, leading LIME to classify the patient as having osteoporosis. Despite this misclassification, the model demonstrated satisfactory accuracy, even with a limited dataset. We believe this machine learning model holds promise as a predictive tool and could eventually be incorporated into osteoporosis treatment guidelines.

Limitations

In this study, we faced several limitations during the collection of patient data. Some data could not be obtained due to equipment-related reasons in hospitals, leading to missing data in the patient population. Additionally, some patients with actual osteoporosis may have been overlooked during hospitalization due to the absence of relevant symptoms, as some nonspecialist doctors may not have been highly attuned to diagnosing osteoporosis. When clinicians inquired about patients' personal histories, such as smoking, some patients may not have admitted to smoking or may have casually denied it during questioning, potentially influencing the research results. The relatively small amount of data collected in this study may have introduced bias. This study involved binary classification and was limited by the use of machine learning methods, without exploring deep learning algorithms or multiclassification methods.

Conclusion

In summary, among the collected patient information, age, height, weight, alkaline phosphatase activity, and osteocalcin were significantly associated with osteoporosis. After multiple evaluations and selection of the most suitable algorithm for prediction, the LGBM emerged as the optimal choice. This study contributes to enhancing the diagnostic capabilities of nonspecialist doctors in identifying osteoporosis and will be valuable for healthcare institutions lacking expertise in diagnosing osteoporosis.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12911-025-02943-7.

Supplementary Material 1

Acknowledgements
Not applicable.

Author contributions
X.G. and P.R. developed the idea. G.H. and W.Z. and wrote an initial draft of the manuscript. Y.L.W., K.C., Y.Z.W., Y.S., W.S. and L.L. were responsible for data collection. X.G. and P.R. wrote the final version of the manuscript. All authors read and approved the final manuscript.

Funding
The National Natural Science Foundation of China (No.81873999, No.82072446).

Data availability
The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
This research had been ethically approved by Huazhong University of Science and Technology Tongji Medical College Medical Ethics Committee (NO. S1060). Informed consent was obtained from all individual participants included in the study. All methods were performed in accordance with the Declarations of Helsinki.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Orthopaedics, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430030, China
[2]Shenzhen Health Development Research and Data Management Center, Shenzhen 518028, China
[3]Department of Orthopedics, Wuhan No. 1 Hospital, Wuhan Integrated TCM & Western Medicine Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
[4]Department of Orthopaedics, Shandong Provincial Hospital, Cheeloo College of Medicine, Shandong University, Jinan, China

Received: 28 August 2023 / Accepted: 20 February 2025
Published online: 11 March 2025

References

1. Zheng LM, Zhuang ZK, Li YX, Shi TS, Fu K, Yan WJ, et al. Bone targeting anti-oxidative nano-iron oxide for treating postmenopausal osteoporosis. Bioact Mater. 2022;14:250–61.
2. Löffler MT, Kallweit M, Niederreiter E, Baum T, Makowski MR, Zimmer C, et al. Epidemiology and reporting of osteoporotic vertebral fractures in patients with long-term hospital records based on routine clinical CT imaging. Osteoporos Int. 2022;33(3):685–94.
3. Pinto D, Alshahrani M, Chapurlat R, Chevalley T, Dennison E, Camargos BM, et al. The global approach to rehabilitation following an osteoporotic fragility fracture: A review of the rehabilitation working group of the international osteoporosis foundation (IOF) committee of scientific advisors. Osteoporos Int. 2022;33(3):527–40.
4. Muschitz C, Hummer M, Grillari J, Hlava A, Birner AH, Hemetsberger M, et al. Epidemiology and economic burden of fragility fractures in Austria. Osteoporos Int. 2022;33(3):637–47.
5. Howlader KC, Satu MS, Awal MA, Islam MR, Islam SMS, Quinn JMW et al. Machine learning models for classification and identification of significant attributes to detect type 2 diabetes. Health Inf Sci Syst. 2022;10(1).
6. Alkady W, ElBahnasy K, Leiva V, Gad W. Classifying COVID-19 based on amino acids encoding with machine learning algorithms. Chemometr Intell Lab. 2022;224.
7. Peng W, Yang Y, Lu H, Shi H, Jiang L, Liao X, et al. Network Pharmacology combines machine learning, molecular simulation dynamics and experimental validation to explore the mechanism of Acetylbinankadsurin A in the treatment of liver fibrosis. J Ethnopharmacol. 2024;323:117682.
8. Wu YW, Cheng MC, Huang S, Pei ZX, Zuo YL, Liu JX et al. Recent advances of deep learning for computational histopathology: principles and applications. Cancers. 2022;14(5).
9. Vukovic DB, Romanyuk K, Ivashchenko S, Grigorieva EM. Are CDS spreads predictable during the Covid-19 pandemic? Forecasting based on SVM, GMDH, LSTM and Markov switching autoregression. Expert Syst Appl. 2022;194:116553.
10. Sun MS, Sun WY, Zhao XT, Li ZQ, Dalbeth N, Ji AC et al. A machine learning-assisted model for renal urate underexcretion with genetic and clinical variables among Chinese men with gout. Arthritis Res Ther. 2022;24(1).
11. Shen J, Liu FS, Xu M, Fu LP, Dong ZH, Wu JC. Decision support analysis for risk identification and control of patients affected by COVID-19 based on bayesian networks. Expert Syst Appl. 2022;196.
12. Islam MR, Nahiduzzaman M. Complex features extraction with deep learning model for the detection of COVID19 from CT scan images using ensemble based machine learning approach. Expert Syst Appl. 2022;195.
13. Haak BW, Brands X, Davids M, Peters-Sengers H, Kullberg RFJ, van Houdt R, et al. Bacterial and viral respiratory tract microbiota and host characteristics in adults with lower respiratory tract infections: A Case-Control study. Clin Infect Dis. 2022;74(5):776–84.
14. Wenhao L, Ren L, Tonghua W, Xiaoqian S, Xiaodong W, Guojie H, et al. Spatio-temporal variation in soil thermal conductivity during the freeze-thaw period in the permafrost of the Qinghai-Tibet plateau in 1980–2020. Sci Total Environ. 2024;913:169654.
15. Makhmutova M, Kainkaryam R, Ferreira M, Min J, Jaggi M, Clay I. Predicting changes in depression severity using the PSYCHE-D (Prediction of severity Change-Depression) model involving Person-Generated health data: longitudinal Case-Control observational study. Jmir Mhealth Uhealth. 2022;10(3).
16. Zhang J, Liu ZZ, Chen RB, Ma QW, Lyu Q, Fu SH, et al. A MALDI-TOF mass spectrometry-based haemoglobin chain quantification method for rapid screen of thalassaemia. Ann Med. 2022;54(1):293–301.
17. Huang R, Cheng Z, Jin XY, Yu XM, Yu JH, Guo YP, et al. Usefulness of four surrogate indexes of insulin resistance in middle-aged population in Hefei, China. Ann Med. 2022;54(1):622–32.
18. Li X, Wang Z, Zhao W, Shi R, Zhu Y, Pan H, Wang D. Machine learning algorithm for predict the in-hospital mortality in critically ill patients with congestive heart failure combined with chronic kidney disease. Ren Fail. 2024;46(1):2315298.
19. Tabassum S, Abedin N, Rahman MM, Rahman MM, Ahmed MT, Islam R, et al. An online cursive handwritten medical words recognition system for busy Doctors in developing countries for ensuring efficient healthcare service delivery. Sci Rep. 2022;12(1):3601.
20. Fudickar S, Nustede EJ, Dreyer E, Bornhorst J, Mask R-CNN, Based C. Elegans detection with a DIY microscope. Biosens (Basel). 2021;11(8). Epub 2021/08/27.
21. Liu YC, Cheng HY, Chang TH, Ho TW, Liu TC, Yen TY, et al. Evaluation of the need for intensive care in children with pneumonia: machine learning approach. JMIR Med Inf. 2022;10(1):e28934.
22. Guo R, Li B, Zeng ZL, Jiang X, Zhang D, Xie TY et al. Thoracolumbar kyphosis in postmenopausal osteoporosis patients without vertebral compression fractures. Ann Transl Med. 2022;10(2).
23. Li YF, Wang QY, Xu LL, Yue C, Hu L, Ding N, et al. Development of a nomogram for predicting very low bone mineral density (T-Scores <-3) in the Chinese population. Int J Gen Med. 2022;15:1121–30.
24. Liu S, Pang Q, Guan W, Yu F, Wang O, Li M et al. Association of serum osteocalcin with bone microarchitecture and muscle mass in Beijing community-dwelling postmenopausal women. Endocrine. 2024.
25. Li S, Chen B, Chen H, Hua Z, Shao Y, Yin H, et al. Analysis of potential genetic biomarkers and molecular mechanism of smoking-related postmenopausal osteoporosis using weighted gene co-expression network analysis and machine learning. PLoS ONE. 2021;16(9):e0257343.
26. Hsu CY, Huang CY, Hsieh CH, Chien PC, Chen CC, Hou SY et al. Regular exercise and Weight-Control behavior are protective factors against osteoporosis for general population: A propensity Score-Matched analysis from Taiwan biobank participants. Nutrients. 2022;14(3).
27. Ren H, Ma X, Shao Y, Han J, Yang M, Wang Q. Correlation between serum miR-154-5p and osteocalcin in males and postmenopausal females of type 2 diabetes with different urinary albumin creatinine ratios. Front Endocrinol (Lausanne). 2019;10:542.
28. Li J, Li Y, Li S, Lu Y, Rai P. Relationship between polymorphisms and mutations at rs7125942 and rs3736228 of LRP5 gene and bone metabolism in postmenopausal women. J Orthop Surg Res. 2024;19(1):104.

Publisher's note