Classification and Prediction of Spinal Disease Based on the SMOTE-RFE-XGBoost Model

Biao Zhang [1] , **Xinyan Dong** [2] , **Yuwei Hu** [2] , **Xuchu Jiang** [Corresp., 2] , **Gongchi Li** [3]

[1] School of Computer Science, Liaocheng University, LIAOCHENG, SHANDONG, CHINA

[2] School of Statistics and Mathematics, Zhongnan University of Economics and Law, WUHAN, HUBEI, CHINA

[3] Union Hospital Affiliated to Tongji Medical College, Huazhong University of Science and Technology, WUHAN, HUBEI, CHINA

Corresponding Author: Xuchu Jiang
Email address: xuchujiang@zuel.edu.cn

Spinal diseases are killers that cause long-term disturbance to people with complex and diverse symptoms and may cause other conditions. At present, the diagnosis and treatment of the main diseases mainly depend on the professional level and clinical experience of doctors, which is a breakthrough problem in the field of medicine. This article proposes the SMOTE-RFE-XGBoost model, which takes the physical angle of human bone as the research index for feature selection and classification model construction to predict spinal diseases. The research process is as follows: two groups of people with normal and abnormal spine conditions are taken as the research objects of this paper, and the synthetic minority oversampling technique (SMOTE) algorithm is used to address category imbalance. Three methods, least absolute shrinkage and selection operator (LASSO) , tree-based feature selection , and recursive feature elimination (RFE) , are used for feature selection. Logistic regression (LR) , support vector machine (SVM) , parsimonious Bayes, decision tree (DT), random forest (RF) , gradient boosting tree (GBT), extreme gradient boosting (XGBoost) , and ridge regression models are used to classify the samples, construct single classification models and combine classification models and rank the feature importance. According to the accuracy and mean square error (MSE) values, the SMOTE-RFE-XGBoost combined model has the best classification, with accuracy, MSE and F1 values of 97.56%, 0.1111 and 0.8696, respectively. The importance of four indicators, lumbar slippage, cervical tilt, pelvic radius and pelvic tilt, was higher.

1   **Classification and Prediction of Spinal Disease Based on the**

2   **SMOTE-RFE-XGBoost Model**

3   Biao Zhang[1], Xinyan Dong[2], Yuwei Hu[2], Xuchu Jiang[2*], Gongchi Li [3]

4   [1]School of Computer Science, Liaocheng University, Liaocheng, 252059, China

5   [2]School of Statistics and Mathematics, Zhongnan University of Economics and Law, Wuhan,

6   430073, China

7   [3]Union Hospital Affiliated to Tongji Medical College, Huazhong University of Science and

8   Technology, Wuhan, 430022, China

9   *Correspondence: xuchujiang@zuel.edu.cn

11   **Abstract:** Spinal diseases are killers that cause long-term disturbance to people with complex

12   and diverse symptoms and may cause other conditions. At present, the diagnosis and treatment of

13   the main diseases mainly depend on the professional level and clinical experience of doctors,

14   which is a breakthrough problem in the field of medicine. This article proposes the SMOTE-

15   RFE-XGBoost model, which takes the physical angle of human bone as the research index for

16   feature selection and classification model construction to predict spinal diseases. The research

17   process is as follows: two groups of people with normal and abnormal spine conditions are taken

18   as the research objects of this paper, and the synthetic minority oversampling technique (SMOTE)

19   algorithm is used to address category imbalance. Three methods, least absolute shrinkage and

20   selection operator (LASSO), tree-based feature selection, and recursive feature elimination

21   (RFE), are used for feature selection. Logistic regression (LR), support vector machine (SVM),

22   parsimonious Bayes, decision tree (DT), random forest (RF), gradient boosting tree (GBT),

23   extreme gradient boosting (XGBoost), and ridge regression models are used to classify the

24   samples, construct single classification models and combine classification models and rank the

25   feature importance. According to the accuracy and mean square error (MSE) values, the

26   SMOTE-RFE-XGBoost combined model has the best classification, with accuracy, MSE and F1

27   values of 97.56%, 0.1111 and 0.8696, respectively. The importance of four indicators, lumbar

28   slippage, cervical tilt, pelvic radius and pelvic tilt, was higher.

30   **Keywords** Spinal diseases; Feature selection; XGBoost; Machine learning; Classification

31   prediction

1 Introduction

Spinal diseases are one of the most common diseases in modern society and are also listed as one of the top 10 diseases affecting health rated by the World Health Organization. According to the Guidelines for the Diagnosis, Treatment and Rehabilitation of Cervical Spondylosis 2019, the prevalence of cervical spondylosis in China is approximately 3.8%-17.6%, and nearly 150 million people in China suffer from cervical spondylosis. A study published in the Lancet showed that of all disease types, cervical spine pain cost the Chinese people the 9th most healthy life expectancy in 2017, up from the 21st in 1990. A survey by the Professional Committee for the Prevention and Treatment of Cervical and Low Back Diseases of the China Association for the Promotion of Healthcare International Exchange showed that the prevalence of cervical spondylosis among young adults aged 20-40 years was as high as 59.1%.

Spinal diseases are caused by pathological changes in the bones, intervertebral discs, ligaments, and muscles of the spine, which in turn compress and stimulate the spinal cord, spinal nerves, blood vessels and vegetative nerves, resulting in complex and diverse symptoms (Byrne et al, 2000) as well as degenerative and infectious diseases of the spine, leading to low back pain and spinal pain or disability and paralysis in more severe cases. At the same time, spinal conditions are also directly or indirectly related to nerve and organ functions in the human body, which can lead to a series of complex diseases. At present, diagnosis and treatment mainly depend on the professionalism and clinical experience of doctors, which places a heavy burden on the problem of low efficiency. Furthermore, the uneven distribution of medical resources in China and the disparity in diagnostic results have made it almost impossible for patients with spinal diseases to seek timely medical care or have accurate diagnosis results, resulting in delayed diagnosis.

In recent years, machine learning and deep learning have played a huge role in many industries. As clinical data always have high dimensionality and large sample sizes, various algorithms can be used as analytical tools to assist in clinical diagnosis. With the advent of the era of healthcare based on big data, artificial intelligence and machine learning technologies can contribute to the rapid and accurate diagnosis of spinal diseases (D'Angelo et al, 2022) (Cabitza et al, 2018) and assist physicians in preoperative planning and postoperative outcome prediction, which could help to improve diagnostic efficiency and reduce the burden on medical staff and the rate of misdiagnosis.

63

1.1 Related Works

1.1.1 An Empirical Study on Spinal Problems in Primary and Middle School Students

In recent years, spinal disorders have become increasingly prevalent at a younger age, and many adolescents, even children, suffer from the pain and distress caused by spinal problems. In addition, the prevalence of spinal disorders among adolescents is still increasing due to factors such as incorrect sitting posture, excessive fatigue caused by long-term maintenance of a specific posture, and frequent use of electronic devices.

Kadhim M et al. (2018) selected a total of 36,728 school students aged 6-18 years in 192 primary and secondary schools in Yuzhong County, Gansu Province, and concluded that the prevalence of spinal abnormalities among primary and secondary school students in Yuzhong County, Gansu Province, was 1.26%. *LI Shangle et al. (2021)* monitored 1902 primary and secondary school students in Lhasa, Shigatse, Nagqu, Chamdo, Shannan city and Ali area and obtained a prevalence of 1.26% for spinal curvature abnormalities in primary and secondary school students. *QI Deyun et al. (2021)* adopted a stratified cluster random sampling method and selected 1,884 primary and secondary school students from grade 4 to grade 3 in Hongkou District for screening of spinal curvature abnormalities and a questionnaire survey of influencing factors. The results showed that the detection rate of spinal curvature abnormalities in primary and secondary school students in Hongkou District was 7.2%.

82

1.1.2 Research on the Application of Different Learning Algorithms in the Diagnosis of Spinal Diseases

A new research direction in the field of artificial intelligence is the cross integration of artificial intelligence technology (machine learning and neural network models) and spinal surgery. At present, many scholars have applied machine learning, deep learning, and computer technology to the diagnosis of spinal diseases and have achieved good results.

In terms of regression analysis methods, *Attiah M et al. (2019)* divided 160 study subjects into four age groups, measured a series of spinal sagittal imaging parameters, analyzed data variance and correlation by one-way ANOVA, the least significant difference method, and the Person test, and identified spinal sagittal imaging parameters that were independently correlated with age by linear regression analysis.

94   In the machine learning algorithm, *Zhu J et al. (2022)* used 348 patients with ankylosing
95   spondylitis as the study subjects, used LASSO, random forest and support vector machine
96   recursive feature elimination to screen feature variables and build prediction models, and finally
97   obtained better prediction results with AUCs of 0.878 and 0.823 on the training and validation
98   sets, respectively. *Müller D et al. (2022)* used 8374 patients who had undergone surgery for
99   degenerative disorders of the spine as the research subjects and performed feature selection and
100   model construction by LASSO and ridge regression methods to finally reach the goal of
101   developing a parsimonious model to predict degenerative thoracic, lumbar, or cervical spine. The
102   aim was to develop a parsimonious model to predict multidimensional outcomes in patients
103   undergoing surgery for thoracic, lumbar or cervical degenerative disorders. Based on orthopedic
104   medical data, *Li S et al. (2020)* constructed and designed a platform for an orthopedic auxiliary
105   diagnosis classification prediction model based on the extreme gradient boosting (XGBoost)
106   algorithm to further realize the auxiliary diagnosis of orthopedic diseases. *Wang S et al. (2018)*
107   proposed combining diffusion tensor imaging (DTI) metrics with machine learning algorithms to
108   accurately classify controls and spinal cervical spondylosis (CSCS). The support vector machine
109   (SVM) classifier produced 95.73% accuracy, 93.41% sensitivity, and 98.64% specificity, which
110   showed significant classification performance advantages.
111   *Shen D et al. (2017)* worked on high-performance deep learning algorithms for medical
112   image processing and analysis, proposed a novel end-to-end multitask structure correlation
113   learning network (MMCL-Net) for the simultaneous detection, segmentation, and classification
114   of three spinal structures (intervertebral disc, vertebrae and neural foramina), and locally
115   optimized the model to achieve a more stable dynamic equilibrium state. *Hu B et al. (2018)* used
116   a long-term short-term memory deep learning network to identify people with chronic low back
117   pain based on human balance and body swing performance in the standing test with an accuracy
118   of 97.2% and a recall rate of 97.2%. *Jamaludin A et al. (2017)* used lumbar spine magnetic
119   resonance imaging (MRI) images as input and classified different tasks by a convolutional neural
120   network (CNN) model with a model accuracy up to 95.6%. *Pedersen C F et al. (2020)*
121   systematically analyzed the accuracy of different models for predicting the postoperative
122   efficacy of lumbar disc herniation. Compared with the traditional machine learning model, the
123   model based on deep learning can better predict the postoperative efficacy of patients.

124   For the dataset, *Raihan-Al-Masud M et al. (2020)* focused on the application of machine
125 learning algorithms for predicting spinal abnormalities. As a data preprocessing step, univariate
126 feature selection as a filter-based feature selection and principal component analysis (PCA) as a
127 feature extraction algorithm are considered. Several machine learning approaches, namely,
128 support vector machine (SVM), logistic regression (LR), and bagging ensemble methods, are
129 considered for the diagnosis of spinal abnormalities.

130

131 **1.1.3 The application of ET and RFE in feature selection**

132   In high-dimensional datasets, feature selection is a key step in extracting important features,
133 and it is also the basis for subsequent modeling. Tree-based feature selection and recursive
134 feature elimination (RFE) are common methods for feature selection that are widely used in
135 different fields for different studies and have achieved good results.

136   In the medical field, *Hu Y et al. (2021)* used preoperative magnetic resonance images of
137 patients with epithelial ovarian tumors. Image features are extracted from the three-dimensional
138 region of interest manually sketched on the axial T2 weighted imaging (T2WI) image. Four
139 feature selection methods and seven machine learning classifiers are combined in pairs. The
140 results show that the RFE-KNN model combined with RFE and the K-nearest neighbor (KNN)
141 classifier has the best performance. *Zhang Y et al. (2022),* based on the physical examination
142 data of a health examination institution in Urumqi in 2018, used three feature selection methods,
143 RFE, measured resting metabolic rate (MRMR) and least absolute shrinkage and selection
144 operator (LASSO), combined with two model explanatory methods of variable importance and
145 linear interpolation with maximum entropy (LIME), to process the metabolic syndrome risk
146 prediction model before and after modeling. *Li L et al. (2020)* used the gene epitope data of
147 spontaneous premature birth (SPB) as a basis, used SVM-RFE for gene feature selection, and
148 compared it with other machine learning and feature selection methods. SPB biomarkers were
149 comprehensively screened out. *Jun-Tao TAN et al. (2021)* used logistic regression (LR), SVM-
150 RFE and elastic net methods to preliminarily screen the characteristic variables of the main
151 factors of liver cirrhosis complicated with hepatic encephalopathy based on the data of patients
152 with liver cirrhosis with complete medical records. LR and multilayer perceptron (MP) were
153 used as meta-learners to construct a stacked generalization (stacking) heterogeneous integrated
154 classification model. *Gitto S et al.(2022)* studied 158 patients with chondroosseous tumors as

155 research objects. This paper classified and diagnosed atypical cartilaginous tumors (ACTs) and
156 grade II chondrosarcomas of long bones (CLB2) using machine learning based on magnetic
157 resonance imaging (MRI) radiomics, applied LASSO and RFE to determine the feature sets for
158 model training, and balanced the datasets using a synthetic minority oversampling technique
159 (SMOTE). The final model achieved good results with 98% and 80% accuracy in ACT and
160 CLB2, respectively.

161 *Cao W et al. (2021)* extracted the 10 most helpful features for predicting purchase behavior
162 based on the ET algorithm and used logistic regression and the support vector machine algorithm
163 to construct a purchase behavior prediction model. Finally, the above two algorithms were fused
164 based on the method of soft voting and obtained a good prediction effect. *Kurniawan R et al.*
165 *(2022)* used Harris hawks optimization (HHO) and support vector regression (HHO-SVR) to
166 build a prediction model for ozone concentration in 14 partitions of JABODETABEK. Recursive
167 feature elimination and support vector regression (RFE-SVR) were used to select the important
168 predictors, the HHO-SVR method and support vector regression (SVR) were used to establish
169 the prediction model, the HHO algorithm was used to optimize the values of their parameters,
170 and the final HHO-SVR model obtained a better conclusion. *Xinyu Wang et al. (2021)* took dam
171 displacement as the research object and safety monitoring data as the research basis and
172 proposed a prediction model for the dam displacement of tailings dams based on recursive
173 feature elimination, random forest and limit gradient enhancement and compared it with
174 prediction models such as XGBoost, LSTM neural networks, BP neural networks and SVR. The
175 results showed that the RFE-RF-XGBoost model had an average relative error of 3.93%, which
176 was lower than that of the XGBoost model.

177 This section first describes the results of an empirical investigation of spine problems in
178 Chinese primary and secondary school students. The results of different algorithms, such as
179 regression analysis, machine learning algorithms, and deep learning algorithms related to the
180 diagnosis of spinal disorders, are also presented separately. Since ET and RFE are often used
181 when dealing with high-level datasets and extracting their important features, the first part also
182 presents the application of the combined model derived from ET and RFE for feature selection.

183

184 **1.2 Research Framework**

185   The physical angle index of human bone can comprehensively reflect the health status of
186   the spine. A small change in a certain index may cause complex chain reactions of other
187   indicators. This study aims to use physical indicators that are relatively unaffected by the cascade
188   effect and then perform feature selection to make the selected features more representative.
189   Furthermore, a machine learning classifier is used to develop a feature selection-classification
190   combination model, and the best performing model is comprehensively compared for spinal
191   health diagnosis. Finally, there is a category imbalance problem in the original sample used in
192   the article, which is the key to the superiority of feature selection and model construction results.
193   In the data processing stage, dealing with the category imbalance by the SMOTE algorithm is
194   one of the focus parts of this article. The research framework of this paper is shown in Fig. 1.
195

2 Research Method

197   According to Fig. 1, the article conducts the empirical analysis part with three main sections:
198   data preprocessing, feature selection and model construction. The important step in the data
199   preprocessing section is the category imbalance treatment using the SMOTE algorithm; the
200   feature selection section uses three models, LASSO, ET and RFE, and the model construction
201   section uses eight different models. To highlight the significant part of the paper, only the
202   SMOTE algorithm, XGBoost model, ET model, and RFE model are described in detail below.
203

2.1 Principle of the SMOTE-ENN Algorithm
205   The category imbalance problem refers to the situation where the number of training
206   samples of different categories in a classification task varies greatly. *Nitesh V. Chawla et al.*
207   *(2022)* proposed the SMOTE algorithm in 2002, which is based on the principle of filtering
208   samples x_i and x_j and the corresponding random numbers $0 < \lambda < 1$ from a small set of samples,
209   constructing new samples $x_n = x_i + \lambda(x_j - x_i)$ by the relationship between the two samples, which
210   means that the minority samples are analyzed and new samples are added to the dataset by
211   artificial synthesis based on the minority samples. The SMOTE algorithm is essentially an
212   oversampling; it does not sample on the data space but in the feature space, so its accuracy will
213   be higher than the traditional sampling method.
214   In the SMOTE algorithm, the neighborhood space is determined by the K-nearest neighbor
215   method, as shown in Fig. 2, assuming that the main sample x has K=5 nearest neighbor samples
216   on the feature space, from which a nearest neighbor sample x' is randomly selected, and then a
217   new sample x^{new} is generated at a random position on the connection line between the main

218   sample x and the nearest neighbor sample x'. The synthesis of the new sample x^{new} is calculated
219   as shown in Eq. (1) is shown. where x is the selected master sample, x' is the randomly selected
220   nearest neighbor sample of x, x^{new} is the synthesized new sample, and rand(0,1) denotes the
221   random number on the generated (0,1).

$$\mathrm{x}^{\mathrm{new}} = x + rand(0,1) \times \left(\mathrm{x}^{'} - x\right) \tag{1}$$

222   The pseudocode of Algorithm 1 of the SMOTE algorithm is shown:

223   **Algorithm 1 *SMOTE***

---

224   **Input:** Training set $S = \{(x_i,y_i), i = 1,2,...,N, y_i \in \{+,-\}\}$; Sample size of majority classes N^{-},

225   Sample size of minority classes N^{+}, $N^{+} + N^{-} = N$; Imbalance ratio $IR = \frac{N^{-}}{N^{+}}$; Sample rate SR;

226   Proximity parameter K.

---

227   **Output:** The training set after oversampling $S = \{(x_i,y_i), i = 1,2,...,N + N^{+} \times SR, y_i \in \{+,-\}\}$

228   1:   Remove all majority class and minority class samples from the training set S to form
229   a majority class training sample set S^{-} and minority classes of training sample sets
230   S^{+}.
231   2:   Set the newly generated sample set S^{New} to empty
232   3: $for\ i = 1:N^{+} \times SR$
233   4:   Randomly choose a number from $[1,N^{+}]$, find the corresponding   sample x in
234   S^{+}.
235   5:   find the K neighbor sample of the main sample x from S^{+}, and set it into the
236   neighbor sample group S^{Near}.
237   6:   Randomly choose a number from $[1,K]$,find the corresponding main neighbor
238   sample $\mathrm{x}^{'}$ from S^{New}.
239   7:   Calculate the new minority class sample: $\mathrm{x}^{new} = x + rand \times (x^{'} - x)$, in which $rand \in$
240          $[0,1]$.
241   8:   Add x^{new} to S^{New}: $S^{New} = S^{New} \cup \mathrm{x}^{new}$
242   9:   Set the neighbor sample group S^{Near}
243   10:   end for
244   return training set after oversampling $S^{'} = S^{-} \cup S^{New}$

245

246   **2.2 The XGBoost Principle**

247   The XGBoost algorithm belongs to the time-series supervised learning model, which is an
248   optimization and enhancement of the gradient boosting tree, by combining several learners to
249   produce a boosting algorithm with strong learners. The loss function of XGBoost includes its
250   own loss and regularization parts, and the addition of the second-order Taylor expansion of the
251   error part and the L_1 and L_2 regularization terms makes it more accurate and has stronger
252   generalization ability. The loss function of XGBoost is shown in Eq. (2) and Eq. (3).

$$L_t = \sum_{i=1}^{J} \left[G_{tj}\omega_{tj} + \frac{1}{2}(H_{tj} + \lambda)\omega_{tj}^2 \right] + \gamma J \tag{2}$$

$$G_{tj} = \sum_{x_i \in R_{tj}} g_{ti}, \quad H_{tj} = \sum_{x_i \in R_{tj}} h_{ti} \tag{3}$$

253 where L_t, which is denoted as the loss function of XGBoost, x_i is the input samples, and G_{tj} and
254 H_{tj} denote the sum of the first-order derivatives and second-order derivatives of all input samples
255 for the tth decision tree mapping to the leaf node j, respectively. J is the number of leaf nodes, γ
256 indicates the difficulty of the node cut, λ is the regularization factor of L_2, g_{ti} and h_{ti} are the first-
257 order and second-order derivatives of the ith sample at the tth weak learner, respectively, and ω_{tj}
258 is the optimal value of the jth leaf node. Finally, the loss function is minimized to obtain the
259 optimal solution ω_{tj} for all J leaf node regions and each leaf node region for the tth decision tree
260 optimum.

261

262 **2.3 Principle of Tree-based Feature Selection**

263 According to the random forest algorithm, the importance of each attribute can be calculated
264 through tree model training, and the value of importance can help us select important features.
265 This study uses the ET algorithm, which directly uses random features and random threshold
266 division, and the shape and difference of each decision tree will be increasingly large and
267 random.

268

269 **2.4 Principle of recursive feature elimination (RFE)**

270 The recursive elimination feature method uses a base model for multiple rounds of training.
271 After each round of training, the features of several weight coefficients are eliminated, and then
272 the next round of training is carried out based on the new feature set. It uses the model accuracy
273 to identify which attributes (and attribute combinations) contribute the most to the prediction of
274 target attributes.

275

276 **2.5 Model Evaluation Metrics**

277 In this paper, the accuracy, recall, F1 value, area under the curve (AUC) value and mean
278 square error (MSE) were used to evaluate the performance of different models. MSE is used to
279 measure the difference between the predicted value and the real value of the regression task,
280 which is the mean of the square sum of the error between the predicted value and the real value,
281 which is calculated as Eq. (4).

$$MSE = \frac{1}{n}\sum_{i=1}^{m}\omega_i(y_i - \hat{y}_i)^2 \tag{4}$$

282   where y_i is the real value, $\hat{y}_i$ is the predicted value, $\omega_i > 0$, and n is the number of samples. The

283   smaller the MSE is, the smaller the difference between the predicted value and the real value.

284   Therefore, the smaller the MSE is, the better the model.

285   Other indicators are calculated based on the confusion matrix. The confusion matrix is a

286   standard format that represents the accuracy evaluation. Each column represents the prediction

287   category, and each row represents the true attribution category of the data. According to the

288   confusion matrix, various indicators can be calculated. The evaluation index and its meaning are

289   shown in Table 1, and the higher the value is, the better the model.

290   This section revolves around the theoretical structure of the research approach of this paper,

291   introducing for the first time a research framework for spinal health diagnosis by feature

292   selection combined with a physical indicator model selected by machine learning. The principles

293   of the SMOTE algorithm, XGBoost, and tree-based feature selection and RFE algorithms

294   employed in the model are presented. In addition, a total of five evaluation metrics, accuracy,

295   recall, F1 value, AUC value, and mean square error (MSE), were adopted as the evaluation

296   metrics of the model.

3 Empirical Analysis

3.1 Data Sources

299   In this paper, the dataset is from the exploration case of dichotomous classification of back

300   pain symptoms on the Kaggle website, which is a dichotomous classification problem to detect

301   whether the spine of a person is healthy by collecting physical data from the human spine and

302   pelvis and other parts. The dataset has 310 sets of sample observations and 13 attributes, of

303   which 12 attributes are numerical variables and serve as independent variables and 1 attribute is

304   a categorical variable and serves as the dependent variable, with the characteristics described in

305   Table 2.

306

3.2 Data Preprocessing

3.2.1 Abnormal Value Processing

309   According to the box diagram, histogram, density curve, etc., there are extreme points in the

310   data, among which x_1 (incidence of pelvic fractures), x_3 (lumbar lordosis angle), x_4 (sacral

311   inclination angle) and x_6 (lumbar spondylolisthesis) have more obvious extreme points. Based on

312   the results of basic descriptive statistical analysis, the maximum incidence of pelvic fractures is

313   129.83% > 100%, which is not consistent with the actual situation; there is a large gap between

314   the maximum values of the lumbar kyphosis angle and sacral inclination angle and their average

315   level and 75% quantile; the maximum value of lumbar spondylolisthesis is obviously separated

316   from the overall level, which can be judged as an abnormal value. Taking the upper and lower

317   boundaries of the box diagram as the critical point of the abnormal value, the abnormal values of

318   the six characteristics of x_1 (incidence of pelvic fractures), x_2 (pelvic tilt), x_3 (lumbar lordosis

319   angle), x_4 (sacral inclination angle), x_5 (pelvic radius) and x_6 (lumbar spondylolisthesis) are

320   identified and eliminated. Among the above 6 features, the data values larger than the upper edge

321   and smaller than the lower edge of the box line graph are considered outliers, and all samples

322   corresponding to the outliers are considered abnormal samples and are rejected. According to the

323   results, the abnormal values of the above six features are 3, 13, 1, 1, 11 and 10. The new dataset

324   after rejection contained 279 samples, and the box diagram and histogram of the new dataset are

325   shown in Fig. 3 and Fig. 4.

326   In Fig. 3, the horizontal axis is the category to which the collected data belong, and $x_1{\sim}x_{12}$

327   refers to the feature symbol in Table 2. The vertical axis is the distribution of the collected data,

328   and 25 is an interval.

329   In Fig. 4, the horizontal axis is the distribution of the collected data, with 20 as an interval.

330   The vertical axis is the density distribution of the collected data, and the orange line is the

331   density curve of the collected data.

332

3.2.2 Category 0-1 Conversion

334   Because the variable is nonnumerical, it is converted to a 0-1 variable. Among them, the

335   normal spine is marked as 0, and the abnormal spine is marked as 1.

336

3.2.3 Min-Max Standardized Processing

338   The dataset was divided into two parts: independent variables and labels. Data

339   standardization (Min-Max) was applied to the independent variables to restrict the data range to

340   [0,1], which eliminated the errors caused by the different magnitudes and transformed all the data

341   into positive values. The *standardization* is used as Eq. (5).

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{5}$$

342   where x is the original value of a feature, $min(x)$ is the minimum of x, $\max(x)$ is the maximum of

343   x and x' is the normalized value of x.

344

345   **3.2.4 Class Imbalance Processing based on SMOTE**

346   After excluding the abnormal values, the dataset showed 99 samples with normal spines and

347   180 samples with abnormal spines, with a large difference in the number of 0-1 samples.

348   Therefore, the SMOTE algorithm was used to process the samples with class imbalance.

349   Specifically, the principle is for a minority class sample. Using the K nearest neighbor method

350   (the k value needs to be specified in advance), the distance xi is obtained. The nearest k minority

351   samples, According to the principle of the SMOTE algorithm, the process of synthesizing a new

352   sample will randomly select one of the 5 nearest neighbor samples, multiply the Euclidean

353   distance of the two samples by a random number between (0,1), and determine the exact location

354   of the synthesized sample based on the new distance. For this dataset, the SMOTE function in

355   the imblearn library (a Python library for handling unbalanced data) is called for oversampling,

356   and the default value of 5 is used for the k_neighbors parameter (the number of neighboring

357   samples k). In the original data, normal spines are the minority sample, and abnormal spines are

358   the majority sample. The SMOTE oversampling process was performed for the minority class

359   samples, and no operation was performed for the majority class samples, which finally expanded

360   the minority class samples to 180 cases. The same number of cases as the majority class samples,

361   of which 99 cases were original samples and 81 cases were synthetic samples, were integrated to

362   form the normal spine class samples.

363

364   **3.3 Feature Selection**

365   The features are selected and compressed using three methods: LASSO, ET, and RFE. All

366   three methods are direct selections of the original features without any linear combination or

367   transformation, and the selected features are consistent with the original features. The feature

368   importance matrix is shown in Table 3.

369

370   **3.3.1 LASSO compression**

371   When performing LASSO compression, the LASSOCV function is used in this paper, and

372   after debugging the optimal regularization parameter alpha, the optimal regularization parameter

373   value of 0.001 is selected, and the final output is the mean value of the residuals and the degree

374   of feature importance using 10-fold cross-validation. According to the LASSO compression

375   results, only features with positive values of the weight parameter are selected; that is, features

376   with 0 and negative values of the parameter are excluded. The final feature selection results are

377   x_2 (pelvic tilt), x_6 (lumbar slippage), x_7 (pelvic obliquity), x_8 (direct tilt), and x_{10} (neck tilt).

378

3.3.2 ET method

380   The importance values of the features under the ET method performed normally. According

381   to the results, 6 features with an importance value higher than 0.06 are selected: x_1 (pelvic

382   fracture incidence), x_2 (pelvic tilt), x_3 (lumbar anterior convexity angle), x_4 (sacral tilt angle), x_5

383   (pelvic radius), and v (lumbar slip degree).

384

3.3.3 RFE method

386   The SVM-RFE algorithm is used to test the optimal number of features, as shown in Fig. 5.

387   When the number of features is 2, 7, 10 and 11, the overall contribution of features reaches the

388   maximum, which is 0.83056. When the number of features is 2, a sparse solution is obtained, and

389   this solution is not desirable. Therefore, the minimum number of features except 2, namely, 7

390   features, is selected. After setting the number parameter to 7, the SVM-RFE algorithm is still

391   used for feature selection, and the value of 1 in Table 3 is the selected feature. The other non1

392   values are the unselected features, which are x_1 (pelvic fracture incidence), x_2 (pelvic tilt), x_3

393   (lumbar anterior convexity angle), x_4 (sacral tilt angle), x_5 (pelvic radius), x_6 (lumbar slippage),

394   and x_{10} (cervical tilt).

395   In Fig. 5, the horizontal coordinate indicates the number of retained features, and the vertical

396   coordinate indicates the overall contribution of the features. Using the average of the SVM

397   model accuracy to represent the contribution level, the higher the accuracy, the greater the

398   overall contribution of the features.

399   According to the above results and analysis, there are three feasible results for feature

400   selection: LASSO compresses the features to 5, selects six features based on tree-based feature

401 selection, and selects seven features based on recursive feature elimination. These three feature

402 selection methods and the selection results are used to construct different classification models.

403

404 **3.4 Model Construction and Algorithm Evaluation**

405 Model construction is divided into two parts. The first part constructs classification models

406 using original features without feature selection and constructs logistic regression, SVM, NB,

407 decision tree (DT) random forest (RF), gradient boosted tree (GBT), XGBoost, and ridge

408 regression models, with a total of eight classification models. The second part is to build the

409 combination model of feature selection-classification; that is, the three methods of feature

410 selection are combined with the seven classification models except the ridge regression model,

411 and 21 different combination models are obtained. In the above two parts, 29 different

412 classification models and combination models were obtained. Five evaluation indexes were

413 compared: accuracy, recall, F1 value, AUC value and MSE. The results of the evaluation metrics

414 are shown in Table 4.

415 In the single-model part, there are 8 models, and the precision of all 8 models is greater than

416 0.8, MSE is less than 0.3, recall, F1 value, and AUC are slightly inferior, but the overall level of

417 the models is good, and the differences between models are not obvious. Among them, the

418 accuracies of SVM, RF, XGBoost, and ridge regression are all greater than 0.9, and the accuracy

419 of ridge regression is the highest at 0.9286, while the AUC and MSE values of RF and XGBoost

420 reach the maximum and minimum at 0.8623 and 0.1389, respectively.

421 In the combined model section, there are 21 models, and the overall model performance

422 under different feature selection methods differed to a lesser extent. However, there are

423 significant differences between individual model performances. Under the LASSO method, the

424 overall accuracy of the models is slightly better than that of the single model, and the recall, F1

425 value, AUC value, and MSE value are inferior to those of the single model. The accuracy of

426 LASSO-DT is the lowest at 0.7955, and the accuracy of the remaining six models is higher than

427 0.87, among which LASSO-SVM and LASSONB have the highest accuracy at 0.9444. Under

428 the ET method, the accuracy of the models is higher than 0.86. The model accuracy is greater

429 than 0.86, in which ET-XGBoost has the highest accuracy of 0.9744, and its MSE value and ET-

430 RF both reach the minimum of 0.1296. Under the RFE method, the overall performance of the

431 model is better, and the best performance of the other four indexes except the recall is in this part.

432   The model accuracy is greater than 0.85, in which RFE-XGBoost has the highest accuracy. The

433   accuracy of RFE-XGBoost is the highest, reaching 0.9756, and its MSE value and ET-SVM are

434   both the smallest, at 0.1111.

435   In summary, the accuracy of the SMOTE-RFE-XGBoost model is the highest in all models

436   (97.56%), the MSE value is the lowest in all models (0.1111), the AUC value is the second

437   highest in all models (0.8834), it is only slightly lower than the top in all models (0.8844), and

438   the F1 value is the third highest in all models (0.8696). Thus, the optimal model of this study is

439   the SMOTE-RFE-XGBoost model.

440

441   **3.5 Feature importance ranking**

442   According to the obtained optimal model SMOTE-RFE-XGBoost, the importance ranking of

443   the features is performed by applying the model, and the features involved in the ranking are

444   processed by feature selection instead of importance analysis for all features. The results of the

445   importance analysis are shown in Fig. 6. According to the results, the importance of lumbar

446   slippage was the highest, and the importance of sacral tilt angle, pelvic radius, and lumbar

447   anterior convexity angle were the next highest, while the importance of lumbar slippage was

448   much higher than that of other features. Therefore, this index of lumbar slippage needs extra

449   attention during clinical diagnosis and model construction.

450   This section revolves around an empirical analysis to detect whether the spine of a person is

451   healthy by using physical data collected by the kaggle website on the human spine and pelvis,

452   among other areas. The dataset contains 310 sets of sample observations and 13 attributes. After

453   preprocessing the data, the data were divided into training and testing sets. The features were

454   selected and compressed using three methods: LASSO, tree-based feature selection, and RFE.

455   The first part constructs a classification model using the original features, and the second part

456   constructs a combined feature selection-classification model. Finally, by comparing the accuracy

457   of each model, the optimal model for this study is obtained: the SMOTE-RFE-XGBoost model.

458   **4 Conclusion and Future Work**

459   This paper deeply explored category imbalance processing, feature selection, and machine

460   learning model combinations to construct a more quasilateral diagnostic model for spinal

461   diseases. For category imbalance processing, the SMOTE algorithm proposed by *Nitesh V.*

462   *Chawla et al. (2002)* is used for oversampling in this paper. Since the dataset used in this paper is

463 a public dataset, *Md. Raihan-Al-Masud et al. (2020)* used this dataset for machine learning

464 classification exploration, including basic data preprocessing, dataset expansion, feature

465 selection, and model construction. Among them, the data preprocessing did not involve category

466 inequality processing, and none of the final obtained accuracies reached 90%. In contrast, the

467 accuracy of the optimal model finally obtained in this paper reached 97.56%, and the

468 classification accuracy was significantly improved. Therefore, the SMOTE category imbalance

469 treatment has an extremely critical impact on model construction, model performance and

470 prediction capability. In terms of feature selection, this paper used the Lasso, ETs and RFE

471 algorithms to directly screen features and did not use PCA for linear transformation, which to a

472 certain extent reduced the computational complexity of the model and retained the most original

473 information of the data. In terms of model construction, all feature selection algorithms are

474 modeled in this paper to select the best feature selection scheme; at the same time, models

475 without feature selection are set up for control to explore the differences in the effects produced

476 by feature selection. From Table 4, we can see that the models with feature selection are better

477 than those without feature selection, and then we compare the classification models constructed

478 by different feature selection algorithms horizontally.

479 The results of this paper show that the combined SMOTE-RFE-XGBoost model has the

480 best classification prediction with an accuracy of 97.56%, which improves the accuracy by 5.89%

481 and reduces the MSE value by 20% compared to the XGBoost single model. The SMOTE

482 category imbalance processing algorithm and RFE feature selection improve the model accuracy

483 to some extent, so the performance of the model is significantly improved when combined with

484 the appropriate classification model. In practical applications, the SMOTE-RFE-XGBoost model

485 can improve the accuracy and reliability of spinal disease diagnosis results to a certain extent and

486 can be used as an aid for the clinical diagnosis of spinal diseases. The dataset used in this study is

487 small in both feature dimensions and sample size. Especially in the era of big data healthcare,

488 where there are extremely high-dimensional features and large sample sizes, machine learning

489 and deep learning can better reflect their strong learning ability in big data. Therefore, more

490 studies and surveys are needed, and more data need to be collected for a deeper study of spinal

491 disorders. Finally, this paper is dedicated to machine learning algorithms, while deep learning

492 and artificial intelligence have shown more complex learning mechanisms and better learning

493 capabilities in recent years. It is worthwhile to use relevant algorithms to explore spinal disorders
494 more deeply.

495   This paper studied the physical data of the human spine and pelvis and finally obtained a
496 unimodal prediction model with good results. In fact, the symptoms of spinal diseases are
497 complex and diverse, and it is obvious that the unimodal model lacks conviction to diagnose only
498 from one aspect of the clinical data, and the model is not robust, so the prediction results could
499 be easily affected, even leading to incorrect diagnosis. The bootstrap method *Md. Raihan-Al-*
500 *Masud et al. (2020)* used to expand the dataset is our main research direction in the future.
501 Therefore, first, the original data need to be expanded in terms of feature dimension and sample
502 size, and second, the model needs to be expanded into a multimodal feature model (collect
503 multifaceted and multiform clinical information such as textual information, digital information
504 and image information) and build a more accurate diagnostic model through multimodal fusion
505 technology, which could ensure accurate prediction and make the model more stable and
506 scientific. It helps to make faster and better decisions to assist clinical diagnosis.

507

508 **Author Contributions:** Conceptualization, X.J. and X.D. ; methodology, B.Z. and Y.H. ; formal
509 analysis, G.L. ; data curation, X.J. ; supervision, X.D. and B.Z. ; writing—original draft
510 preparation, X.D. and B.Z. ; writing—review and editing, X.J.

511 **Method:** All methods were carried out in accordance with relevant guidelines and regulations.

512 **Funding:** This research received no external funding.

513 **Data Availability:** The public dataset is available at Kaggle:

514 https://www.kaggle.com/alihussain1993/lower-back-pain-symptoms-datasetlabeled

515 **Conflicts of Interest:** The authors declare that there are no conflicts of interest.

References

517 [1] Byrne T N, Waxman S G, Benzel E C. 2000. Diseases of the spine and spinal cord[M].
518   Oxford University Press.

519 [2] D'Angelo T, Caudo D, Blandino A, et al. 2022. Artificial intelligence, machine learning and
520   deep learning in musculoskeletal imaging: Current applications. Journal of Clinical
521   Ultrasound 50(9): 1414-1431.

522 [3] Cabitza, F., Locoro, A., & Banfi, G. 2018. Machine learning in orthopedics: a literature
523   review. Frontiers in bioengineering and biotechnology, 6, 75.

524   [4] Kadhim M, Lucak T, Schexnayder S, et al. 2020. Current status of scoliosis school screening:
525   targeted screening of underserved populations may be the solution[J]. public health 178: 72-
526   77.

527   [5] LI Shangle, ZHANG Qiang, RONG Xia. 2021. Abnormal spinal curvature in primary and
528   secondary school students in Tibet. Chinese Journal of School Health 42(9): 1384-1387.
529   DOI 10.16835/j.cnki.1000-9817.2021.09.026

530   [6] QI Deyun, LI Liping, JIANG Yanwei, ZHOU Yuan. Abnormal spinal curvature and
531   influencing factors in children and adolescents in Hongkou District, Shanghai. CHINESE
532   JOURNAL OF SCHOOL HEALTH, 2021, 42(3): 444-447, 453. DOI 10.16835/j.cnki.1000-
533   9817.2021.03.030

534   [7] Attiah M, Gaonkar B, Alkhalid Y, et al. 2019. Natural history of the aging spine: a cross-
535   sectional analysis of spinopelvic parameters in the asymptomatic population. Journal of
536   Neurosurgery: Spine 32(1): 63-68.

537   [8] Zhu J, Lu Q, Liang T, et al. 2022. Development and validation of a machine learning-based
538   nomogram for prediction of ankylosing spondylitis. Rheumatology and Therapy 9(5): 1377-
539   1397. DOI 10.1007/s40744-022-00481-6. Epub ahead of print. PMID: 35932360.

540   [9] Müller D, Haschtmann D, Fekete TF, Kleinstück F, Reitmeir R, Loibl M, O'Riordan D,
541   Porchet F, Jeszenszky D, Mannion AF. 2022. Development of a machine-learning based
542   model for predicting multidimensional outcome after surgery for degenerative disorders of
543   the spine. Eur Spine J. 31(8):2125-2136. DOI 10.1007/s00586-022-07306-8. Epub 2022 Jul
544   14. PMID: 35834012.

545   [10] Li S, Zhang X. 2020. Research on orthopedic auxiliary classification and prediction model
546   based on XGBoost algorithm. Neural Computing and Applications 32(7): 1971-1979.

547   [11] Wang S, Hu Y, Shen Y, et al. 2018. Classification of diffusion tensor metrics for the
548   diagnosis of a myelopathic cord using machine learning. International journal of neural
549   systems 28(02): 1750036.

550   [12] Shen D, Wu G, Suk H I. 2017. Deep learning in medical image analysis. Annual review of
551   biomedical engineering 19: 221-248.

552   [13] Hu B, Kim C, Ning X, et al. 2018. Using a deep learning network to recognise low back pain
553   in static standing. Ergonomics 61(10): 1374-1381.

554 [14] Jamaludin A, Lootus M, Kadir T, et al. 2017. ISSLS PRIZE IN BIOENGINEERING
555   SCIENCE 2017: Automation of reading of radiological features from magnetic resonance
556   images (MRIs) of the lumbar spine without human intervention is comparable with an expert
557   radiologist. European spine journal 26(5): 1374-1383.

558 [15] Pedersen C F, Andersen M Ø, Carreon L Y, et al. 2020. Applied machine learning for spine
559   surgeons: predicting outcome for patients undergoing treatment for lumbar disc herniation
560   using PRO data. Global Spine Journal : 2192568220967643.

561 [16] Raihan-Al-Masud M, Mondal M R H. 2020. Data-driven diagnosis of spinal abnormalities
562   using feature selection and machine learning algorithms. Plos one, 15(2): e0228422.

563 [17] Hu Y, Liu Y, Zheng YE, et al. 2021. Application of different radiomics models based on
564   MRI conventional T2WI in preoperative tri-classification of ovarian epithelial tumors.
565   Chinese Journal of Magnetic Resonance Imaging 12(12):34-38,54. DOI:10.12015/issn.1674-
566   8034.2021.12.007

567 [18] Zhang Y, Razbek J, Li D, et al. 2022. Construction of Xinjiang metabolic syndrome risk
568   prediction model based on interpretable models. BMC Public Health 22(1): 1-11.

569 [19] Li L, Liu Z P. 2020. Biomarker discovery for predicting spontaneous preterm birth from
570   gene expression data by regularized logistic regression. Computational and Structural
571   Biotechnology Journal 18: 3434-3446.

572 [20] Jun-Tao TAN, Xiao-Mei XU, Yu-Xin HE, et al. 2021. Construction of prediction model of
573   cirrhosis-related hepatic encephalopathy based on machine learning algorithm. Medical
574   Journal of Chinese People's Liberation Army 46(4): 354-360.

575 [21] Gitto S, Cuocolo R, van Langevelde K, et al. 2022. MRI radiomics-based machine learning
576   classification of atypical cartilaginous tumour and grade II chondrosarcoma of long bones.
577   EBioMedicine 75: 103757.

578 [22] Cao W, Wang K, Gan H, et al. 2021. User online purchase behavior prediction based on
579   fusion model of CatBoost and Logit. Journal of Physics: Conference Series. IOP Publishing
580   2003(1): 012011.

581 [23] Kurniawan R, Setiawan I N, Caraka R E, et al. 2022. Using Harris hawk optimization towar
582   ds support vector regression to ozone prediction. Stochastic Environmental Research and Ris
583   k Assessment 36(2): 429-449.

584   [24] Xinyu Wang, Peng Yang, Jianfei Dai. 2021. Prediction of dam displacement based on RFE-
585   RF-XGBoost. Xinyu Wang, Peng Yang, Jianfei Dai 53(02):60-66.DOI:10.16163/j.cnki.22-1
586   123/n.2021.02.012.

587   [25] Nitesh V. Chawla, Kevin W. Bowyer, Lawrence O. Hall, W. Philip Kegelmeyer. 2002. SMO
588   TE: Synthetic Minority Over-sampling Technique. The Journal of Artificial Intelligence Res
589   earch 16(0).

590   [26] Raihan-Al-Masud, M., & Mondal, M. R. H. (2020). Data-driven diagnosis of spinal abnorma
591   lities using feature selection and machine learning algorithms. Plos one, 15(2), e0228422.

Figure 1

Figure 1 Research framework
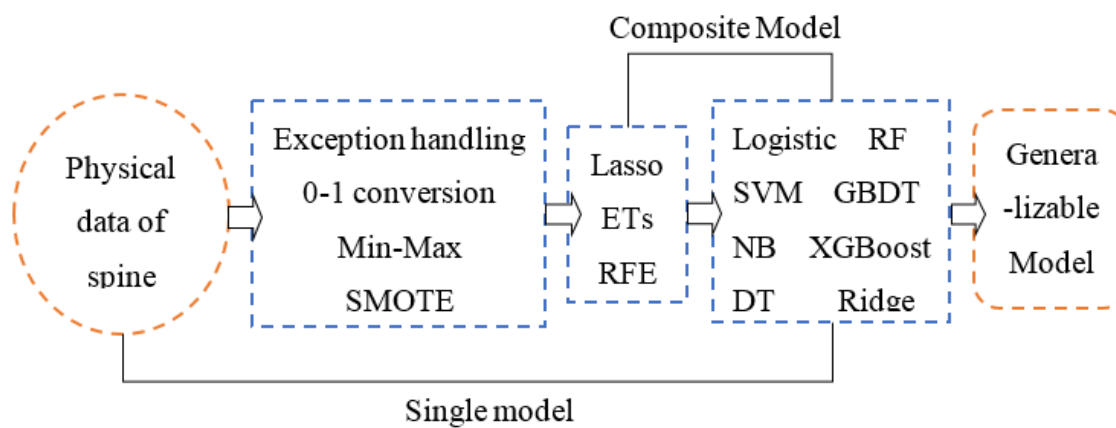
Figure 2

Figure 2 SMOTE algorithm mechanism
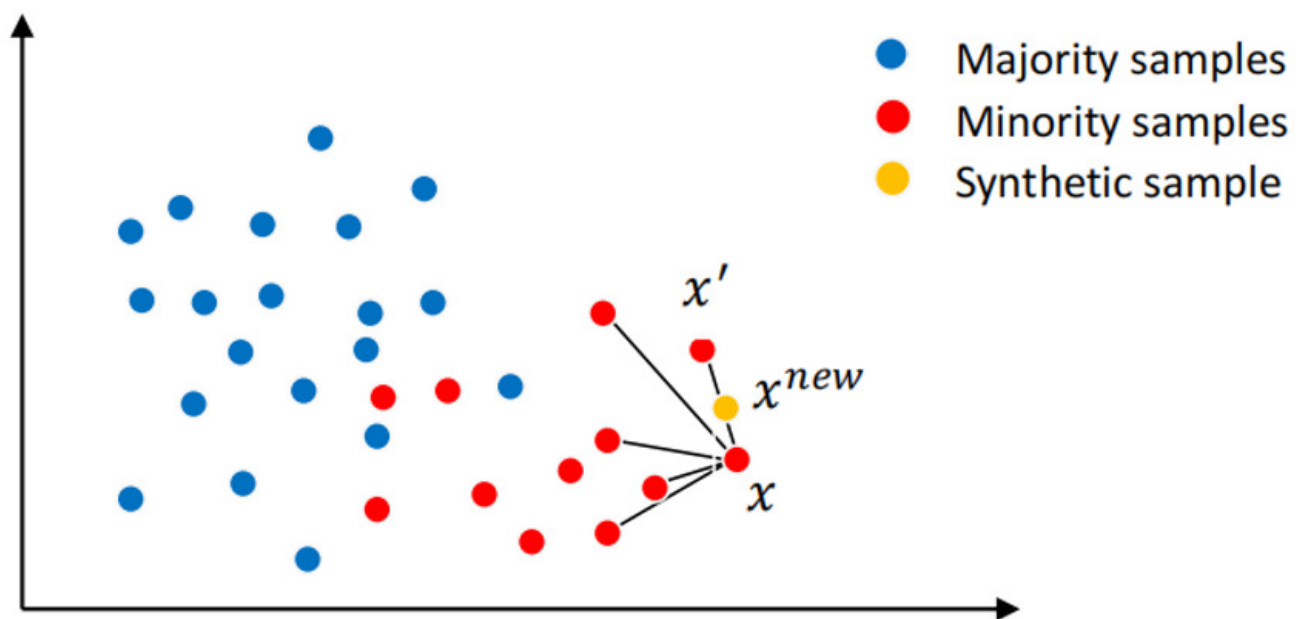
Figure 2 SMOTE algorithm mechanism

Figure 3

Figure 3 Feature boxplot
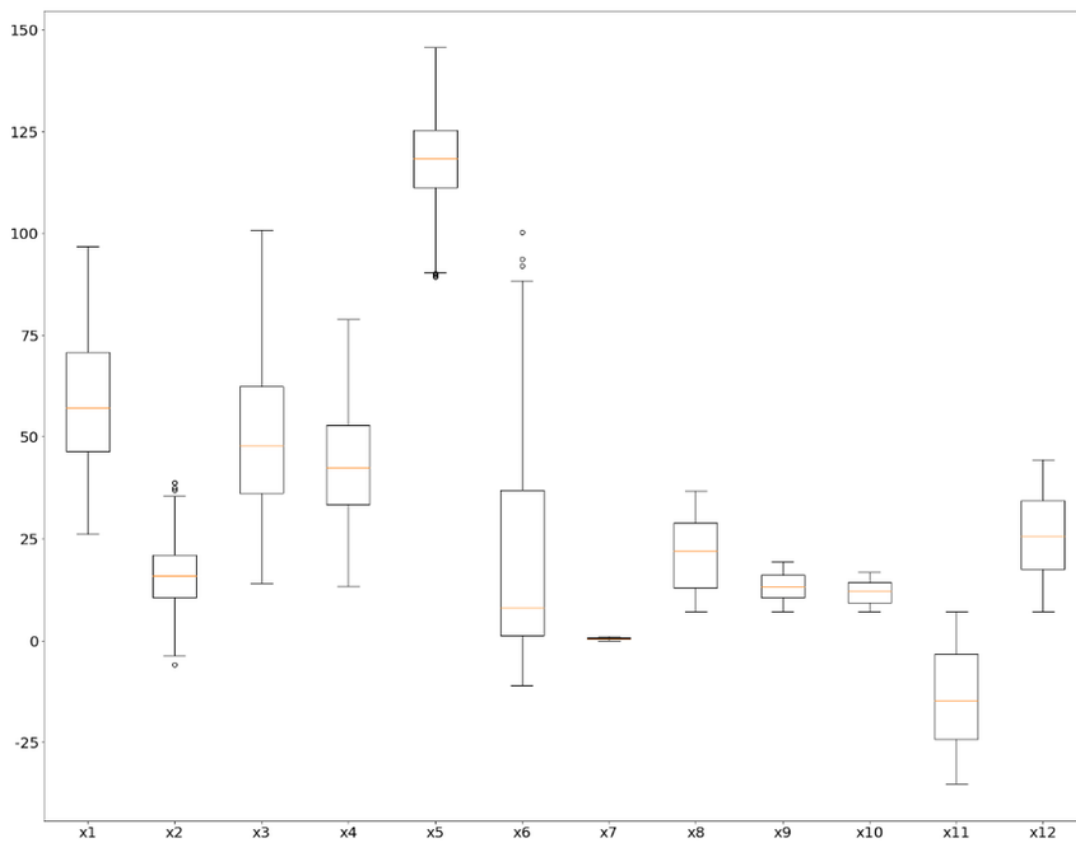
Figure 4

Figure 4 Characteristic histogram and density curve
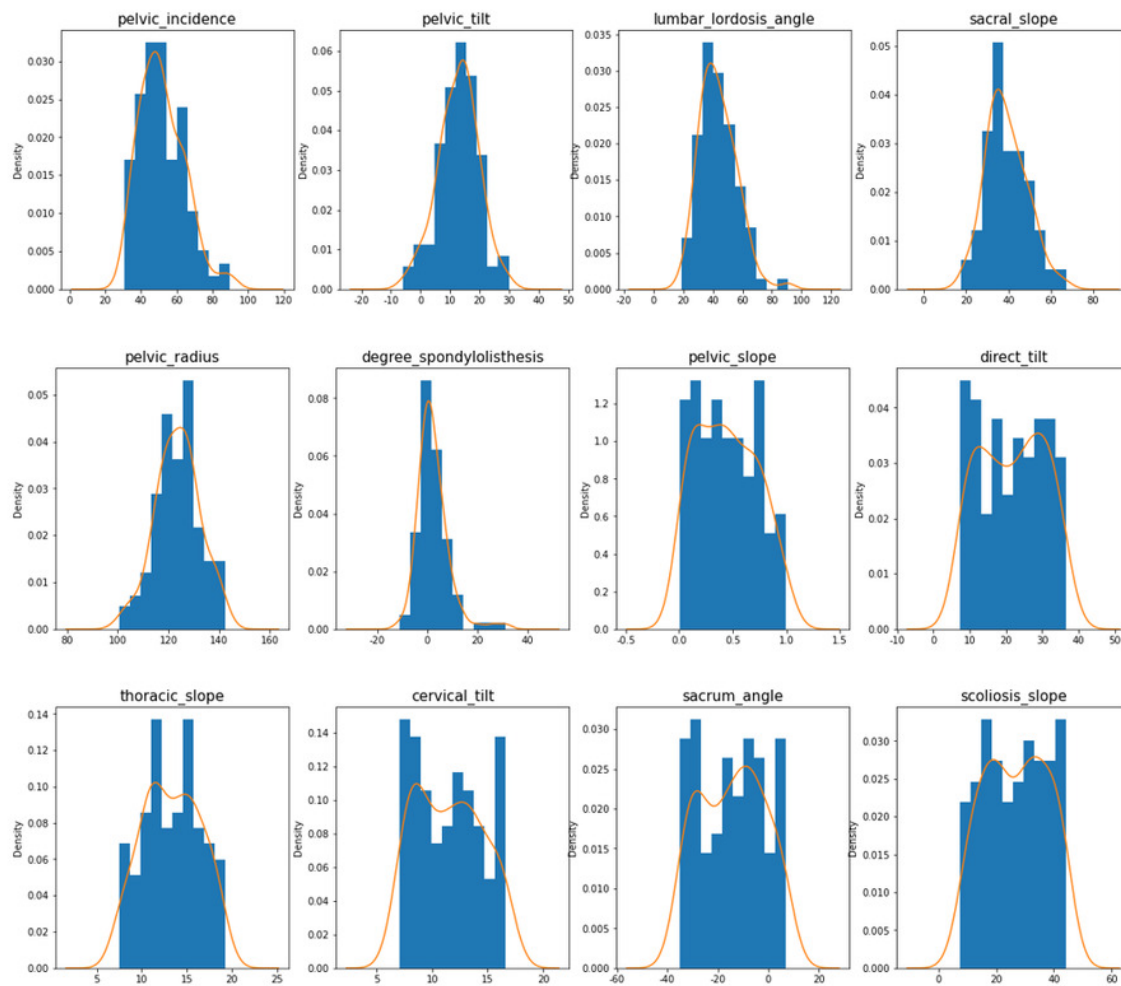
Figure 5

Figure 5 Optimal Feature Number Selection under SVM-RFE
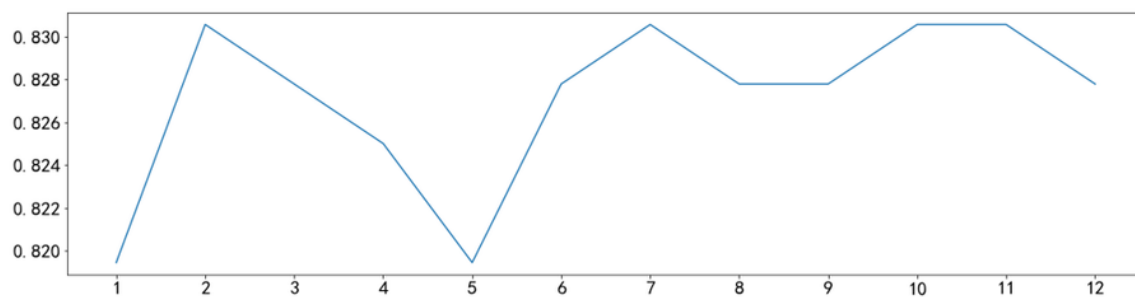
Figure 6

Figure 6 SMOTE-RFE-XGBoost feature importance ranking
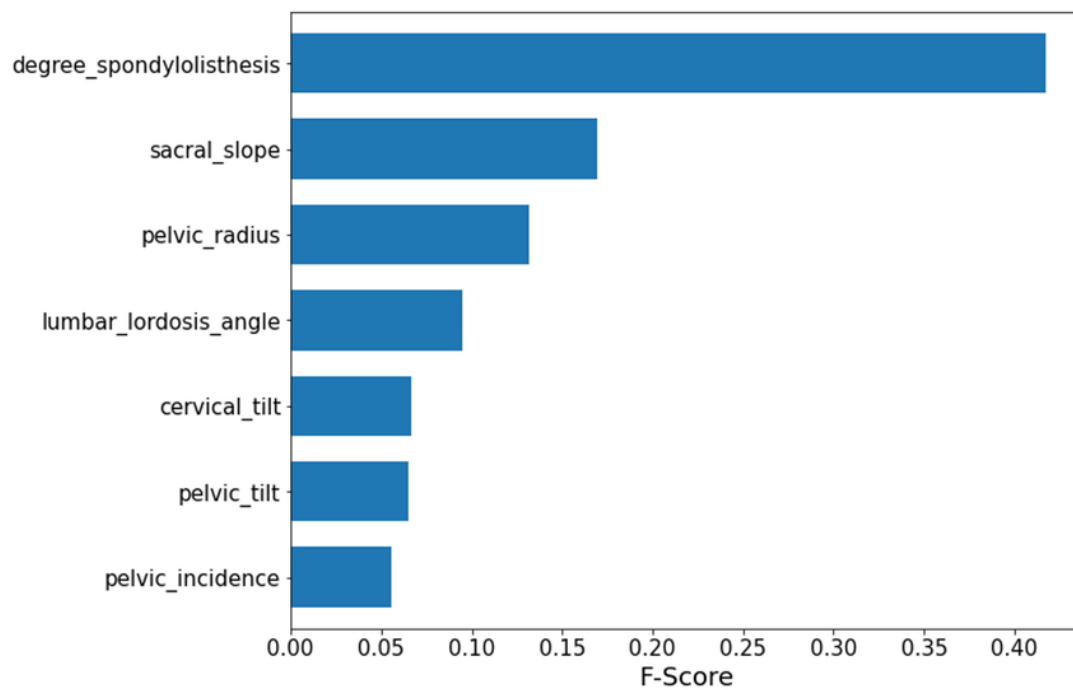
Table 1 Evaluation index

Table 1 Evaluation index

Indicators	Meaning
Accuracy	The proportion of the number of paired samples to the total number of samples
Recall rate	The proportion of the number correctly divided into positive cases to all positive cases
$F1$ value	Comprehensive index of recall and precision, measuring the equilibrium point on the curve P_R
AUC value	The area enclosed by ROC curve and horizontal axis is used to measure the generalization performance of the model

Table 2 Characteristic description

Table 2 Characteristic description

Variables	Feature Name	Feature symbol	Feature Type
Independent variable	Pelvic incident	x_1	Continuous type
	Pelvic tilt	x_2	
	Lumbar lordosis angle	x_3	
	Sacral slope	x_4	
	Pelvic radius	x_5	
	Degree spondylolisthesis	x_6	
	Pelvic slope	x_7	
	Direct tilt	x_8	
	Thoracic slope	x_9	
	Cervical tilt	x_{10}	
	Sacrum angle	x_{11}	
	Scoliosis slope	x_{12}	
Dependent variable	class_att	y	Discrete type

Table 3 Feature Importance Matrix

Table 3 Feature Importance Matrix

Variables	LASSO	ETs	RFE
x_1	0.0000	0.0762	1
x_2	0.2166	0.0746	1
x_3	-0.1679	0.0733	1
x_4	-0.5076	0.0785	1
x_5	-0.8561	0.1297	1
x_6	1.5406	0.2854	1
x_7	0.0070	0.0409	6
x_8	0.0024	0.0448	5
x_9	-0.0526	0.0514	2
x_{10}	0.0619	0.0520	1
x_{11}	-0.0602	0.0490	4
x_{12}	-0.0924	0.0441	3

Table 4 Results of model evaluation indexes

Table 4 Results of model evaluation indexes

Type	Models	Accuracy	Recall rate	F1 value	AUC value	MSE Value
Single model	Logistic	0.8605	0.6727	0.7551	0.7798	0.2222
	SVM	0.9111	0.7455	0.8200	0.8350	0.1667
	NB	0.8222	0.6727	0.7400	0.7609	0.2407
	DT	0.8519	**0.8364**	0.8440	0.8427	0.1574
	RF	0.9000	0.8182	0.8571	0.8619	0.1389
	GBT	0.8980	0.8000	0.8462	0.8528	0.1481
	XGBoost	0.9167	0.8000	0.8544	0.8623	0.1389
	RR	0.9286	0.7091	0.8041	0.8262	0.1759
Combi-nation model	LASSO-Logistic	0.9231	0.7059	0.8000	0.8266	0.1667
	LASSO-SVM	0.9444	0.6667	0.7816	0.8158	0.1759
	LASSO-NB	0.9444	0.6667	0.7816	0.8158	0.1759
	LASSO-DT	0.7955	0.6863	0.7368	0.7642	0.2315
	LASSO-RF	0.8780	0.7059	0.7826	0.8091	0.1852
	LASSO-GBT	0.8780	0.7059	0.7826	0.8091	0.1852
	LASSO-XGBoost	0.9250	0.7255	0.8132	0.8364	0.1574
	ETs-Logistic	0.8864	0.7647	0.8211	0.8385	0.1574
	ETs-SVM	0.9091	0.7843	0.8421	0.8571	0.1389
	ETs-NB	0.8780	0.7059	0.7826	0.8091	0.1852
	ETs-DT	0.8723	0.8039	0.8367	0.8493	0.1481
	ETs-RF	0.9111	0.8039	0.8542	0.8669	0.1296
	ETs-GBT	0.8667	0.7647	0.8125	0.8297	0.1667
	ETs-XGBoost	0.9744	0.7451	0.8444	0.8638	0.1296
	RFE-Logistic	0.8864	0.7647	0.8211	0.8385	0.1574
	RFE-SVM	0.9535	0.8039	**0.8723**	**0.8844**	0.1111
	RFE-NB	0.8750	0.6863	0.7692	0.7993	0.1944
	RFE-DT	0.8542	0.8039	0.8283	0.8406	0.1574
	RFE-RF	0.9302	0.7843	0.8511	0.8658	0.1296
	RFE-GBT	0.8864	0.7647	0.8211	0.8385	0.1574
	RFE-XGBoost	**0.9756**	0.7843	0.8696	0.8834	**0.1111**