



OPEN ACCESS

EDITED BY

Gino Seravalle,
University of Milano-Bicocca, Italy

REVIEWED BY

Hotimah Masdan Salim,
Nahdlatul Ulama University of
Surabaya, Indonesia
Sun Ningling,
People's Hospital, Peking
University, China

*CORRESPONDENCE

Wenlong Zhao
cqzhaowl@163.com

SPECIALTY SECTION

This article was submitted to
Hypertension,
a section of the journal
Frontiers in Cardiovascular Medicine

RECEIVED 28 September 2022

ACCEPTED 17 November 2022

PUBLISHED 01 December 2022

CITATION

Wang T, Yan Y, Xiang S, Tan J, Yang C
and Zhao W (2022) A comparative
study of antihypertensive drugs
prediction models for the elderly
based on machine learning algorithms.
Front. Cardiovasc. Med. 9:1056263.
doi: 10.3389/fcvm.2022.1056263

COPYRIGHT

© 2022 Wang, Yan, Xiang, Tan, Yang
and Zhao. This is an open-access
article distributed under the terms of
the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution
or reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

A comparative study of antihypertensive drugs prediction models for the elderly based on machine learning algorithms

Tiantian Wang¹, Yongjie Yan², Shoushu Xiang³, Juntao Tan⁴,
Chen Yang¹ and Wenlong Zhao^{1*}

¹School of Medical Informatics, Chongqing Medical University, Chongqing, China, ²Medical Records and Statistics Office, The Third Affiliated Hospital of Army Medical University, Chongqing, China, ³Medical Records and Statistics Room, Affiliated Banan Hospital of Chongqing Medical University, Chongqing, China, ⁴Operation Management Office, Affiliated Banan Hospital of Chongqing Medical University, Chongqing, China

Background: Globally, blood pressure management strategies were ineffective, and a low percentage of patients receiving hypertension treatment had their blood pressure controlled. In this study, we aimed to build a medication prediction model by correlating patient attributes with medications to help physicians quickly and rationally match appropriate medications.

Methods: We collected clinical data from elderly hypertensive patients during hospitalization and combined statistical methods and machine learning (ML) algorithms to filter out typical indicators. We constructed five ML models to evaluate all datasets using 5-fold cross-validation. Include random forest (RF), support vector machine (SVM), light gradient boosting machine (LightGBM), artificial neural network (ANN), and naive Bayes (NB) models. And the performance of the models was evaluated using the micro-F1 score.

Results: Our experiments showed that by statistical methods and ML algorithms for feature selection, we finally selected Age, SBP, DBP, Lymph, RBC, HCT, MCHC, PLT, AST, TBIL, Cr, UA, Urea, K, Na, Ga, TP, GLU, TC, TG, γ -GT, Gender, HTN CAD, and RI as feature metrics of the models. LightGBM had the best prediction performance with the micro-F1 of 78.45%, which was higher than the other four models.

Conclusion: LightGBM model has good results in predicting antihypertensive medication regimens, and the model can be beneficial in improving the personalization of hypertension treatment.

KEYWORDS

elderly hypertension, ML, LightGBM, personalized medicine, antihypertensive drug

Introduction

Hypertension is one of the most common chronic diseases and a significant risk factor for cardiovascular disease, chronic kidney disease, cognitive impairment, all-cause mortality, and disability (1). Therefore, the elderly population is a priority population of concern. The relationship between the elderly and the prevalence of hypertension has been demonstrated in studies in various countries (2–6). From 1976 to 2017, the majority of hypertension among women and men aged 60 years or older fluctuated from 60 to 75% in a 123-item national health screening survey of 12 high-income nations and reached this high level in the early 2000s and has been reasonably consistent since then (7). The same trend was found in the seventh report of the Joint National Committee (JNC) on Prevention, Detection, Evaluation, and Treatment of Hypertension, with more than 65% of older adults suffering from hypertension (8). At the same time, older patients have a lower control rate than younger patients in hypertension (9). With the acceleration of population aging, the number of elderly hypertension patients is also increasing yearly, causing a heavy burden on the world's public health.

Maintaining appropriate blood pressure is a critical strategy for lowering the incidence of cardiovascular illness, disability, and mortality. Nowadays, the treatment options for hypertension mainly include life interventions and pharmacological treatment. The common antihypertensive drugs can be divided into five categories: angiotensin-converting enzyme inhibitors (ACEI), angiotensin II receptor antagonists (ARB), calcium channel blockers (CCB), β -blockers, and diuretics. According to the biggest research ever conducted on hypertension, between 1990 and 2019, fewer than one-quarter of hypertensive women and one-fifth of hypertensive men could regulate their blood pressure to normal levels with medication (10). The World Health Organization (WHO) and the International Society of Hypertension (ISH) recommended individualized regimens for the treatment of patients with hypertension (11). The JNC recommended employing 12-lead ECG, urinalysis, blood glucose, hematocrit, serum potassium, creatinine, calcium levels, and fasting lipoprotein profiles, which included high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, and triglycerides (8). The importance of individualizing hypertension treatment is further illustrated.

Electronic Health Record (EHR) is the digitization of an individual's health record. It contains patient data collected during clinical care, including diagnostic billing codes, procedure codes, vital signs, laboratory test results, clinical imaging, and physician records. In recent years, the EHR has become a centerpiece of hospital information systems as healthcare information technology continues to evolve and the demand for clinical and record information management

surges. EHR data are used in various applications, including epidemiological and observational studies, safety surveillance and regulatory use, and prospective clinical studies. Examples: Tabesh et al. (12) analyzed shifts in antihypertensive and lipid-lowering medication therapy in individuals with type 2 diabetes between 2006 and 2015 using EHR data. The US Food and Drug Administration (FDA) conducted post-market safety investigations using EHR data from various sources (13). Jernberg et al. (14) used EHR data to improve care and develop treatments for coronary artery disease based on scientific evidence. According to Cowie (15), EHR should improve clinical diagnosis.

The research on ML of hypertension involves the construction of hypertension risk prediction models (16, 17) as well as the risk prediction research of hypertension complicated by other diseases (18, 19). Recently, Gideon et al. applied ML techniques like decision tree and neural network to find out how to treat hypertension in a large group of Maccabi Health Service patients in Israel (20). Their analysis revealed that 17,234 people were treated with a single medicine, much more than the number of patients treated with a combination treatment. The effectiveness of monotherapy was around 44%. Although this success rate was unsatisfactory, it was still more than the success rate of two, three, or four medication combinations. Liu et al. (21) plotted the profile of five widely used antihypertensive medications (Irbesartan, Metoprolol, Felodipine, Amlodipine, Levamlodipine). Cr and Cys were the common crucial factors affecting its drug control rate. In addition, K^+ , FPG, and LDL also impacted Irbesartan; Metoprolol was affected by Age, Urea, and HCT; Felodipine was affected by FPG, Age, and HB; and Amlodipine was affected by FPG, Age, and Urea. Liu et al. (22) found that HCT, HB, CR, Urea, CYS, age, sex, SBP, DBP, PP, and HR were significant and sensitive factors influencing the rate of hypertension management with Metoprolol. These studies might help tailor the administration of effective antihypertensive drugs to the individual patient, but no final models accounting for these variables were built. Our study used EMR data to describe drug-related features from a vast data set. Then, we constructed medication prediction algorithms to optimize individualized medicine usage and give a reference for physicians.

Materials and methods

Data sources and population

The data we collected comes from the MeduCloud data platform of Chongqing Medical University. It brings together 40 million EMR data from seven general hospitals in the western city of Chongqing, China, containing basic patient information, admission records, medical course records, test

and examinations, symptoms and signs, and medication information. The study did not involve human or animal participants, and all data were identified before collection to comply with ethical requirements.

Data were included from 2016.12 to 2021.12 for inpatients aged ≥ 65 years who had registered antihypertensive drugs in the medication information and whose diagnosis of “primary hypertension” was recorded in the discharge report. Although the MeduCloud data platform documented the antihypertensive medicine usage of many outpatients, there was insufficient clinical test data to match the requirements of this research; thus, only inpatients were included.

Data extraction

We chose gender and age at the time of presentation from basic information, blood pressure from admission records, and a history of hypertension (HTN), coronary artery disease (CAD), diabetes, stroke, and renal insufficiency (RI). In the beginning, 47 indicators were chosen from the clinical tests based on the recommendations of specialists. However, certainly expected indications were absent from the majority of samples. In addition to the fact that professionals pay more attention to the essential characteristics during real diagnosis and therapy, we ultimately chose 22 attributes. The 22 attributes were: glucose (GLU), total cholesterol (TC), triglycerides (TG), potassium (K), sodium (Na), calcium (Ca), creatinine (Crea), serum uric acid (UA), serum urea (Urea), aspartate aminotransferase (AST), alanine aminotransferase (ALT), γ -glutamyl transpeptidase (γ -GT), total bilirubin (TBIL), total protein (TP), red blood cell (RBC), hematocrit (HCT) level, mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), lymphocyte (Lymph) count, platelet (PLT) count, and mean platelet volume (MPV). The standard reference ranges for blood pressure and laboratory indices were listed in [Supplementary Table S1](#).

Statistical analysis

We used SPSS 26.0 and Excel 2013 software for statistical analysis. Kolmogorov-Smirnov was used to test data normality. $x \pm s$ expressed the normal distribution, and the comparison between groups was performed by analysis of variance. The non-normal distribution data were described as M (Q1, Q3), and the Kruskal-Wallis test was used to compare groups. The categorical indicators were expressed by rate (%), and the χ^2 -test was used to compare groups. $P < 0.05$ was considered a statistically significant difference. For the algorithm, we relied on Python (version 3.97).

Machine learning algorithms

In this study, we used five ML algorithms for feature selection, namely Pearson correlation coefficient, Lasso, maximum information coefficient (MIC), random forest (RF), and recursive feature elimination (REF). The five ML models were built into an ensemble model according to specific rules. Finally, combined the ensemble model with statistical analysis to obtain the final feature indicators. For a better comparison, values of all attributes were Min-Max normalized to a standard range using the following formula:

Normalized value = (original value – minimized value) / (maximized value – minimized value)

We compared the performance of RF, SVM, LightGBM, ANN, and NB, which are all well-known ML classification algorithms, to find the best model for predicting medication use. To segment the data for each algorithm, we used a 5-fold cross-validation approach, trained the model through the training set each time, and validated the model’s performance on the test set data. Due to the imbalance between the three predicted medication classes in this study, we used micro-F1 to assess the model’s efficacy. Higher values of micro-F1 indicated better model performance, and it can be used for both multi-class classification problems and asymmetrical data. The calculation is as follows (for three categories):

$$\begin{aligned} Recall_m &= \frac{TP1 + TP2 + TP3}{TP1 + TP2 + TP3 + FN1 + FN2 + FN3} \\ Precision_m &= \frac{TP1 + TP2 + TP3}{TP1 + TP2 + TP3 + FP1 + FP2 + FP3} \\ micro - F1 &= 2 \frac{Recall_m \times Precision_m}{Recall_m + Precision_m} \end{aligned}$$

TP_i refers to a true positive of class i ; FP_i refers to a false positive of class i ; TN_i refers to a true negative of class i ; FN_i refers to a false negative of class i .

The flow chart of the data process and analysis is shown in [Figure 1](#).

Results

Brief introduction of the cases selected for the study

We looked up “2016.12-2021.12, age ≥ 65 years, primary hypertension” on the MeduCloud platform to get a sample of 51,895 cases. Firstly, we excluded 14,950 cases that did not register the use of antihypertensive drugs. Secondly, removed 10,884 cases lacked blood pressure and laboratory test records. Finally, deleted 15,043 cases that did not have complete records of 26 clinical indicators. After data cleaning, 11,018 cases were left. In light of the actual

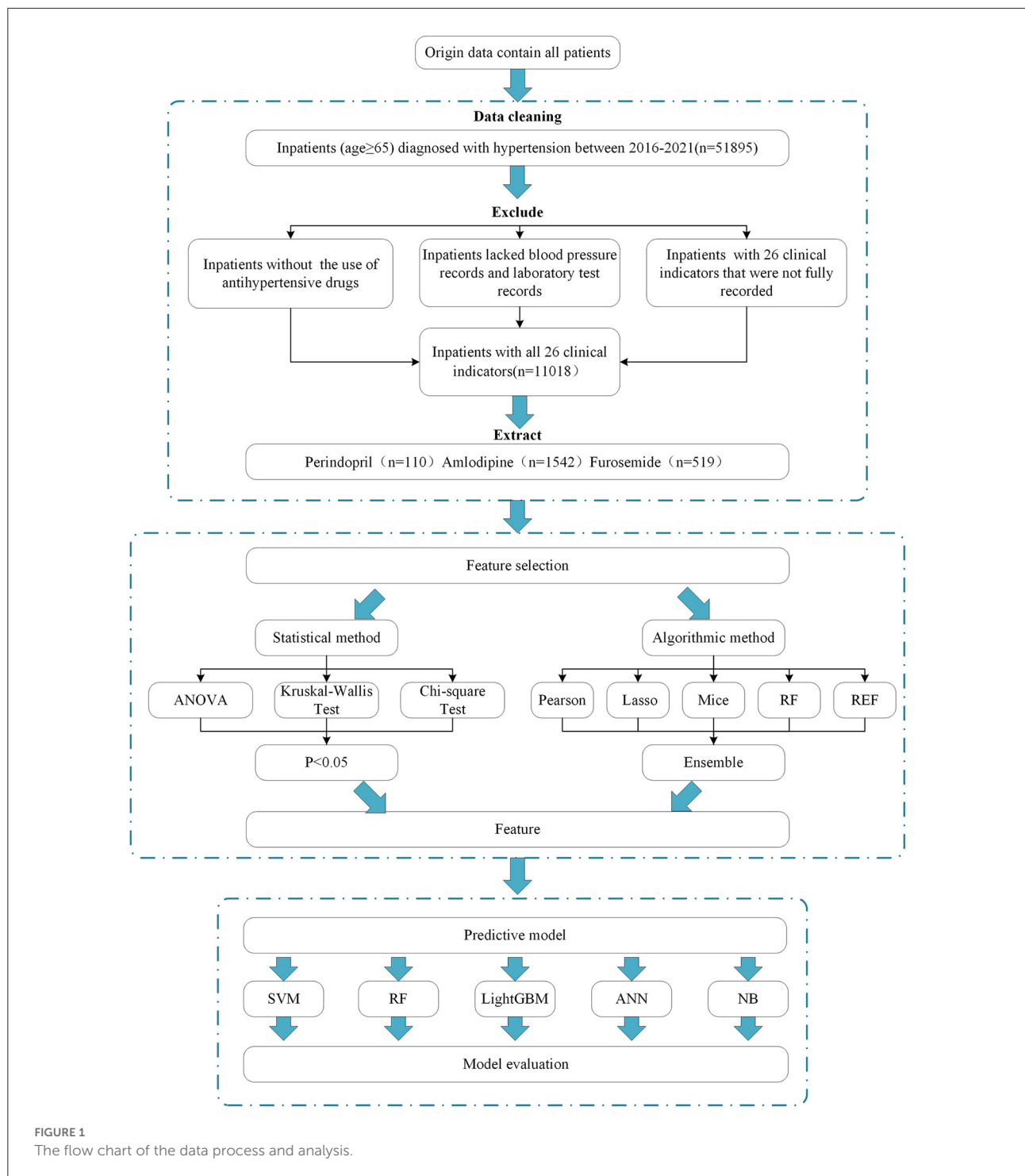
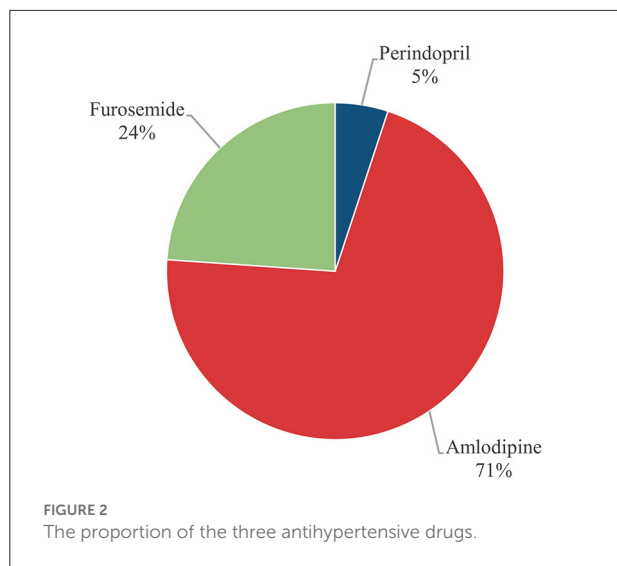


FIGURE 1 The flow chart of the data process and analysis.

situation, we selected three more frequent drugs for studying. They were 110 cases of Perindopril, 1,542 cases of Amlodipine, and 519 cases of Furosemide (Figure 2). Meanwhile, we counted the number of patients (36,945) taking antihypertensive drugs (Supplementary Table S2, Supplementary Figure S1).

Test the difference in the indexes in patients using different kinds of drugs

The results showed that there was a significant difference in age between the groups ($H = 107.33, P = 0.000$), and the age of patients in the group using Furosemide [79.00 (73.00,



84.00] was higher than in the other two groups [74.00 (69.00, 79.00)]. In laboratory tests, there were significant differences in SBP, DBP, Lymph, RBC, HCT, MCHC, PLT, AST, TBIL, Cr, UA, Urea, K, Na, Ga, TP, GLU, TC, TG, and γ -GT ($P < 0.05$, Table 1). There was a significant difference in the proportion of males and females using drugs in the three groups ($\chi^2 = 20.665$, $P = 0.000$). Men use Amlodipine and Furosemide more frequently than women, and in contrast, women use Perindopril more regularly than men. Among previous disease history, there was a statistically significant difference in HTN ($\chi^2 = 31.664$, $P = 0.000$), CAD ($\chi^2 = 24.805$, $P = 0.000$), and RI ($\chi^2 = 11.806$, $P = 0.003$) (Table 1).

ML feature selection

The purpose of feature selection is to pick a set number of feature subsets that provide the least generalization error on the original sample or to decide the smallest feature subset possible under a given generalization error (23). Applying the concept of integration learning, we developed our own rules to combine the outputs of the five algorithms:

1. We recorded the feature ranking obtained by each algorithm. The higher the ranking number, the more influential the feature was.
2. We found the average order of each feature under the five algorithms.
3. We calculated the weight corresponding to the ranking average.

According to the algorithm analysis, Lymph, Urea, HCT, TC, Na, TP, RBC, TBIL, K, SBP, Age, TG, UA, and

AST were the essential features affecting the selection of medication for hypertension. The outcomes of the above feature selection procedures were summed together, with deeper colors indicating more crucial features (Table 2, Figure 3).

Combination of statistical analysis and algorithm results

Combine statistical analysis with ML feature selection, where 1 (dark red) indicates that this feature is considered important by both statistics and algorithms, and 0 (light white) indicates that it is not regarded as relevant by either. If the algorithm does not value a feature, but the statistics do, we have a 0.5 (light red) feature, and vice versa (Figure 4).

In order not to omit the information related to the medication regimen, in the subsequent analysis, we still choose Age, SBP, DBP, Lymph, RBC, HCT, MCHC, PLT, AST, TBIL, Cr, UA, Urea, K, Na, Ga, TP, GLU, TC, TG, γ -GT, Gender, HTN, CAD, RI as feature indicators.

Prediction model performance

On the basis of these indicators, we built five ML models. Table 3 displays the results from the five ML models. We mainly compared the magnitude of the micro-F1 values. The micro-F1 value for each ML algorithm is the average of the five results from the 5-fold cross-validation. LightGBM has the best performance, with a micro-F1 value of 78.4%, significantly higher than the other four ML algorithms (77.7, 73.1, 72.4, and 71.6%, respectively).

We ranked the significance of all variables in the LightGBM model to comprehend the role of each better. The findings revealed that the five most critical characteristics of the model, TP, TBIL, Na, Lymph, and UA, contributed considerably to the prediction outcomes (Figure 5).

We also figured out the saprophytic additivity explanation (SHAP) values of the LightGBM for different types of drugs to find the most important factors that drive model predictions. Figures 6A–C show the significant features of each class. Na, TBIL, and Ca increased, and γ -GT and TC decreased, favoring the classifier to predict Perindopril. Higher ALT and SBP, lower AST, TBIL, and Urea helped the classifier to predict Amlodipine. Higher levels of UA and lower levels of SBP, Lymph, HCT, and Na were beneficial for the classifier to predict Furosemide. Figure 6 depicts other critical characteristics of each medication class. We also visualized the interaction between features (Supplementary Figure S2).

TABLE 1 Test the difference in the indexes in patients using different kinds of drugs.

| Variable | Perindopril | Amlodipine | Furosemide | c2/H/F | P |
|--|-------------------------|-------------------------|-------------------------|--------|--------|
| Male [cases (%)] | 46 (41.80) | 883 (57.30) | 274 (52.80) | 20.67 | <0.001 |
| HTN [cases (%)] | 61 (55.50) | 1,156 (75.00) | 419 (80.70) | 31.66 | <0.001 |
| CAD [cases (%)] | 16 (14.50) | 273 (17.70) | 142 (27.40) | 24.81 | <0.001 |
| RI [cases (%)] | 0 (0.00) | 5 (0.30) | 12 (2.30) | 11.81 | <0.001 |
| Diabetes [cases (%)] | 19 (17.30) | 296 (19.20) | 123 (23.70) | 5.50 | 0.060 |
| Stroke [cases (%)] | 7 (6.40) | 122 (7.90) | 31 (6.00) | 2.31 | 0.320 |
| Age [year, M (Q1, Q3)] | 74.00 (69.00, 79.00) | 74.00 (69.00, 79.00) | 79.00 (73.00, 84.00) | 107.33 | <0.001 |
| SBP [mmHg, M (Q1, Q3)] | 143.50 (128.50, 155.50) | 144.00 (131.00, 158.00) | 134.00 (120.00, 147.00) | 102.81 | <0.001 |
| DBP [mmHg, M (Q1, Q3)] | 81.00 (72.00, 89.25) | 80.00 (72.00, 88.00) | 74.00 (65.00, 84.00) | 67.81 | <0.001 |
| Lymph [$\times 10^9/L$, M (Q1, Q3)] | 1.42 (1.05, 1.76) | 1.41 (1.07, 1.80) | 1.04 (0.73, 1.46) | 158.85 | <0.001 |
| RBC [$\times 10^{12}/L$, M (Q1, Q3)] | 4.33 (4.03, 4.71) | 4.26 (3.91, 4.61) | 3.83 (3.35, 4.30) | 162.97 | <0.001 |
| HCT [% , M (Q1, Q3)] | 40.15 (36.53, 43.23) | 39.00 (36.10, 42.00) | 35.00 (30.20, 39.40) | 157.10 | <0.001 |
| MCV [fl, M (Q1, Q3)] | 92.00 (89.45, 94.58) | 92.10 (88.80, 95.10) | 92.60 (88.60, 96.80) | 5.29 | 0.070 |
| MCH [pg, M (Q1, Q3)] | 30.50 (29.50, 31.60) | 30.50 (29.40, 31.60) | 30.50 (28.80, 32.00) | 0.11 | 0.950 |
| MCHC [g/L, M (Q1, Q3)] | 332.00 (324.75, 338.00) | 332.00 (324.00, 339.00) | 329.00 (319.00, 338.00) | 14.58 | <0.001 |
| PLT [$\times 10^9/L$, M (Q1, Q3)] | 172.00 (127.75, 215.00) | 189.00 (151.75, 232.00) | 169.00 (128.00, 233.00) | 31.16 | <0.001 |
| MPV [fl, M (Q1, Q3)] | 11.10 (9.90, 12.30) | 10.90 (10.00, 11.90) | 10.80 (9.80, 11.80) | 4.82 | 0.090 |
| AST [U/L, M (Q1, Q3)] | 20.00 (17.00, 25.00) | 20.20 (16.8, 25.32) | 23.00 (18.00, 38.00) | 53.82 | <0.001 |
| ALT [U/L, M (Q1, Q3)] | 16.00 (11.35, 22.00) | 16.55 (11.90, 24.70) | 17.00 (11.00, 30.50) | 1.94 | 0.380 |
| TBIL [$\mu\text{mol/L}$, M (Q1, Q3)] | 11.70 (9.18, 14.89) | 10.30 (7.90, 13.45) | 11.70 (7.90, 17.20) | 27.45 | <0.001 |
| Cr [$\mu\text{mol/L}$, M (Q1, Q3)] | 72.15 (58.80, 89.3) | 67.39 (55.90, 81.93) | 82.70 (62.60, 125.30) | 101.52 | <0.001 |
| UA [$\mu\text{mol/L}$, M (Q1, Q3)] | 323.35 (273.93, 412.05) | 325.98 (259.00, 394.40) | 372.50 (280.50, 485.60) | 58.77 | <0.001 |
| Urea [mmol/L, M (Q1, Q3)] | 5.77 (5.05, 7.39) | 5.68 (4.61, 7.13) | 7.33 (5.50, 11.07) | 148.58 | <0.001 |
| K [mmol/L, M (Q1, Q3)] | 4.03 (3.77, 4.24) | 3.93 (3.64, 4.20) | 4.06 (3.66, 4.50) | 30.89 | <0.001 |
| Na [mmol/L, M (Q1, Q3)] | 141.40 (139.78, 143.13) | 141.23 (139.00, 142.95) | 139.50 (136.00, 142.80) | 55.86 | <0.001 |
| Ca [mmol/L, M (Q1, Q3)] | 2.25 (2.19, 2.35) | 2.24 (2.14, 2.33) | 2.19 (2.05, 2.30) | 43.10 | <0.001 |
| TP [g/L, $\bar{x} \pm s$] | 68.94 \pm 6.87 | 68.90 \pm 6.77 | 65.49 \pm 8.28 | 20.99 | <0.001 |
| GLU [mmol/L, M (Q1, Q3)] | 5.71 (4.92, 7.05) | 5.77 (5.03, 7.31) | 6.30 (5.10, 8.36) | 18.93 | <0.001 |
| TC [mmol/L, M (Q1, Q3)] | 4.19 (3.62, 4.90) | 4.37 (3.66, 5.10) | 3.70 (2.92, 4.58) | 123.77 | <0.001 |
| TG [mmol/L, M (Q1, Q3)] | 1.31 (0.97, 1.87) | 1.29 (0.95, 1.75) | 1.09 (0.80, 1.51) | 42.07 | <0.001 |
| r-GT [U/L, M (Q1, Q3)] | 21.78 (15.84, 28.02) | 23.00 (16.68, 36.60) | 33.00 (18.00, 63.00) | 59.89 | <0.001 |

Discussion

Previous research has extensively discussed the significance of medication (24, 25). However, only about a third of persons with hypertension have it under control. The JNC study further supports the need for hypertension patients to get tailored treatment plans (26, 27). This study used SVM, RF, LightGBM, ANN, and NB algorithms to construct a prediction model combining patients' general information, history, and laboratory test findings. The proposed model performed well on the micro-F1 evaluation index and could assist clinicians in predicting medication regimens for elderly hypertension patients.

Amlodipine belongs to CCB, and several international, large-scale clinical investigations have proved its safety and efficacy in middle-aged and older populations. These studies

include VALUE, ALLHAT, and ASCOTA (28–30). It also has a good effect in combination with other antihypertensive drugs. Furosemide, a diuretic that is both cheap and generally well-tolerated by patients, is a crucial tool for lowering blood pressure to normal in older adults with poorly controlled hypertension (31). Perindopril is an ACEI that reduces the risk of cardiovascular events and preserves target organs in elderly hypertensive individuals (32).

Hypertension generally causes kidney, heart, and other disease abnormalities, and antihypertensive medications may alter liver function. As a result, hypertensive patients usually need their blood routine, blood biochemistry, liver and kidney function, cardiac enzymes, and other items tested. Based on the literature reports (33, 34), clinical expertise, and data attenuation (more features, fewer patients fitting

TABLE 2 Results of ML models for feature selection.

| Corr. | Weight | Lasso | Weight | MIC | Weight | RF | Weight | RFE | Weight | Ensemble | Weight |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|----------|--------|
| HCT | 1.00 | Lymph | 1.00 | Urea | 1.00 | Lymph | 1.00 | Lymph | 1.00 | Lymph | 0.97 |
| Urea | 0.79 | Urea | 0.76 | RBC | 0.92 | Urea | 0.89 | K | 1.00 | Urea | 0.92 |
| Lymph | 0.68 | HCT | 0.47 | Lymph | 0.91 | HCT | 0.82 | Ca | 1.00 | HCT | 0.88 |
| Age | 0.61 | TC | 0.38 | HCT | 0.90 | AST | 0.66 | CAD | 1.00 | TC | 0.83 |
| SBP | 0.58 | Na | 0.31 | UA | 0.88 | Na | 0.64 | RI | 1.00 | Na | 0.73 |
| TC | 0.56 | Age | 0.28 | Na | 0.85 | SBP | 0.62 | HTN | 0.96 | TP | 0.64 |
| TP | 0.45 | TG | 0.28 | TC | 0.85 | TC | 0.59 | Stroke | 0.92 | RBC | 0.61 |
| UA | 0.44 | GLU | 0.26 | Cr | 0.84 | RBC | 0.58 | TC | 0.88 | TBIL | 0.60 |
| DBP | 0.36 | TP | 0.24 | TP | 0.76 | TP | 0.54 | Gender | 0.85 | K | 0.60 |
| TBIL | 0.36 | TBIL | 0.19 | K | 0.75 | TBIL | 0.53 | DM | 0.81 | SBP | 0.59 |
| Na | 0.33 | MCV | 0.16 | r-GT | 0.74 | K | 0.47 | TG | 0.77 | Age | 0.57 |
| RBC | 0.30 | SBP | 0.12 | TBIL | 0.73 | r-GT | 0.45 | Urea | 0.73 | TG | 0.55 |
| K | 0.28 | MCHC | 0.10 | AST | 0.68 | UA | 0.44 | HCT | 0.69 | UA | 0.51 |
| AST | 0.26 | MPV | 0.09 | GLU | 0.64 | Age | 0.41 | RBC | 0.65 | AST | 0.50 |
| r-GT | 0.24 | AST | 0.04 | ALT | 0.61 | GLU | 0.37 | MCH | 0.62 | Ca | 0.48 |
| Ca | 0.19 | DBP | 0.03 | SBP | 0.58 | PLT | 0.36 | MCV | 0.58 | GLU | 0.47 |
| TG | 0.14 | ALT | 0.02 | PLT | 0.57 | Ca | 0.36 | Na | 0.54 | r-GT | 0.39 |
| CAD | 0.13 | UA | 0.02 | Ca | 0.52 | TG | 0.35 | MCHC | 0.50 | MCHC | 0.38 |
| MCHC | 0.11 | Cr | 0.01 | TG | 0.51 | MCHC | 0.30 | MPV | 0.46 | MCV | 0.37 |
| RI | 0.11 | r-GT | 0.01 | MCV | 0.46 | ALT | 0.29 | Age | 0.42 | DBP | 0.35 |
| ALT | 0.09 | Gender | 0.00 | MCH | 0.45 | Cr | 0.29 | TBIL | 0.38 | CAD | 0.34 |
| GLU | 0.08 | RBC | 0.00 | DBP | 0.40 | DBP | 0.26 | TP | 0.35 | HTN | 0.32 |
| HTN | 0.08 | MCH | 0.00 | Age | 0.37 | MCV | 0.25 | GLU | 0.31 | RI | 0.32 |
| Cr | 0.07 | PLT | 0.00 | MCHC | 0.36 | MPV | 0.24 | SBP | 0.27 | Cr | 0.31 |
| MPV | 0.05 | K | 0.00 | MPV | 0.16 | MCH | 0.22 | DBP | 0.23 | ALT | 0.30 |
| DM | 0.03 | Ca | 0.00 | HTN | 0.06 | Gender | 0.03 | UA | 0.19 | MPV | 0.28 |
| MCV | 0.02 | HTN | 0.00 | CAD | 0.05 | HTN | 0.03 | AST | 0.15 | MCH | 0.25 |
| PLT | 0.02 | DM | 0.00 | RI | 0.03 | DM | 0.02 | ALT | 0.12 | Gender | 0.23 |
| Stroke | 0.01 | Stroke | 0.00 | Gender | 0.02 | CAD | 0.02 | Cr | 0.08 | PLT | 0.23 |
| Gender | 0.00 | CAD | 0.00 | DM | 0.01 | Stroke | 0.00 | r-GT | 0.04 | DM | 0.23 |
| MCH | 0.00 | RI | 0.00 | Stroke | 0.00 | RI | 0.00 | PLT | 0.00 | Stroke | 0.23 |

all parameters), we chose 22 blood test indicators that are somewhat associated with hypertension. In the blood routine, WBC is an important indicator of the inflammatory response, partly reflecting the control level of blood pressure in elderly hypertensive patients (35). Gagnon et al. (36) proposed that high HCT levels are related to potential risk factors for cardiovascular disease. Yang (34) found that RBC, Hb, HCT, MCHC, RDW-SD, and RDW-CV were all higher in the elderly hypertension group than in the control group ($P < 0.05$), suggesting that secondary erythrocytosis occurs in hypertensive patients and may induce an increase in blood viscosity. The MPV of the hypertensive group was lower than that of the control group ($P < 0.05$), indicating that evaluating PLT and platelet parameters might be employed as hypertension preventive and therapy detection indications

in the elderly. Blood potassium, sodium, liver and kidney function, blood glucose, lipids, and other factors are all examined in blood biochemistry. If the test results show low blood potassium, secondary hypertension is possible. The detection of blood glucose and blood lipids may assist in determining if there are additional risk factors for cardiovascular and cerebrovascular disorders, and the evaluation of liver and kidney function is beneficial for physicians to pick antihypertensive drugs according to the patient's state. The results of these blood test indicators affect the choice of antihypertensive drugs and may guide the treatment of hypertension.

Despite the fact that there are defined clinical guidelines for the management of hypertension, there is still some confusion about how to treat hypertension. Therefore, we

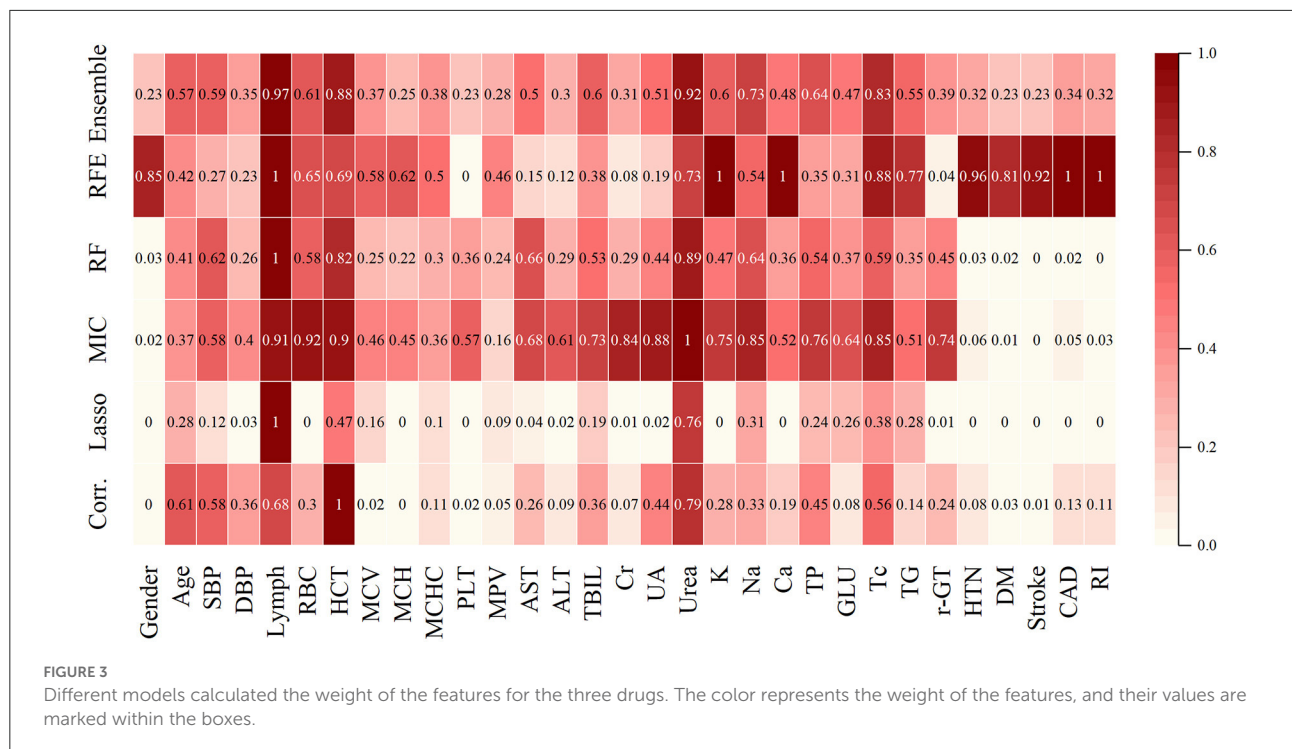


FIGURE 3 Different models calculated the weight of the features for the three drugs. The color represents the weight of the features, and their values are marked within the boxes.

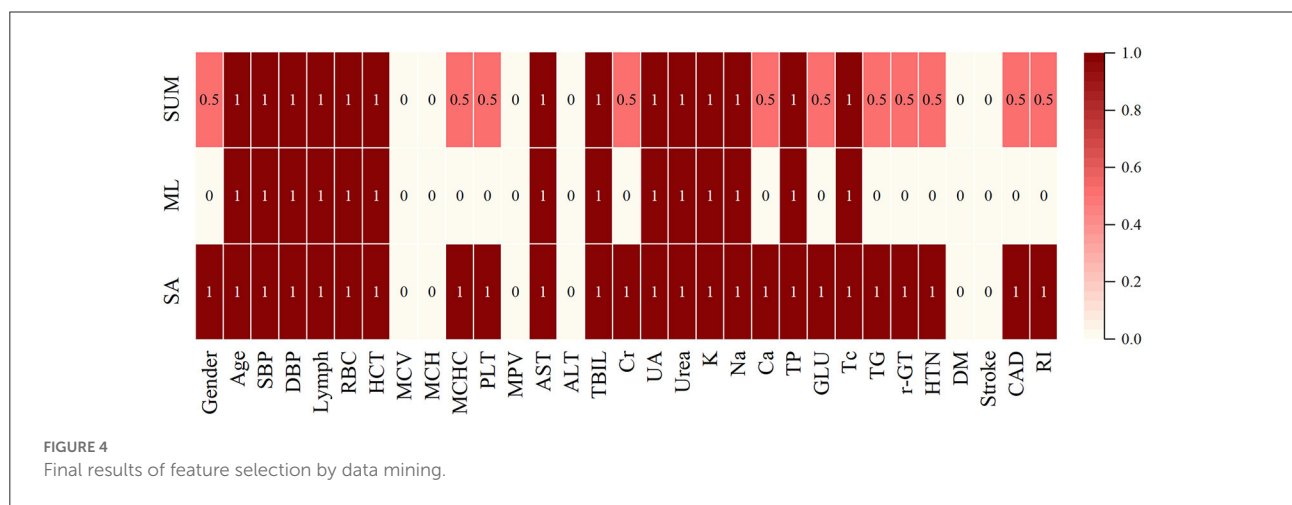


FIGURE 4 Final results of feature selection by data mining.

develop individualized medication by determining which characteristics influence drug selection. Our results showed that Age, SBP, DBP, Lymph, RBC, HCT, MCHC, PLT, AST, TBIL, Cr, UA, Urea, K, Na, Ga, TP, GLU, TC, TG, γ -GT, Gender, HTN, CAD, and RI were the typical indicators of the model. And among them, TP, TBIL, Na, Lymph, and UA were the five main indicators of LightGBM, which critically affected the selection of antihypertensive drugs. Wei et al. (37) discovered that the lower TP levels in elderly hypertensive patients were related to age, DBP, LDH, and HDL-C levels. So, monitoring TP is of great clinical reference value in the assessment and management of hypertension

in the elderly. According to Wang et al. (38), UA and TBIL are essential markers of metabolic abnormalities in hypertensive patients, which are directly related to their development and can reflect the severity of hypertension. Kunutsor et al. (39) found that circulating TBIL levels were independently associated with an increased risk of cardiovascular disease. However, the relationship between TBIL levels in circulating blood and hypertension has not been elucidated. Numerous studies have shown that sodium in the diet influences blood pressure and that cutting down on salt helps bring blood pressure down (40–42). The INTERSALT research, a massive cross-sectional analysis, found

that sodium was strongly linked to increased blood pressure with age (43). Guzik et al. proposed that hypertension enhanced T lymphocyte production in response to tumor necrosis factor-alpha and that T cells may provide a unique therapeutic target for hypertension therapy (44). At the same time, the ratio of neutrophils to lymphocytes is a special inflammatory indication that may be used to predict, treat, and forecast the course of primary hypertension by focusing on inflammation (45). The differences in indicators found between the groups helped direct our selection of antihypertensive drugs.

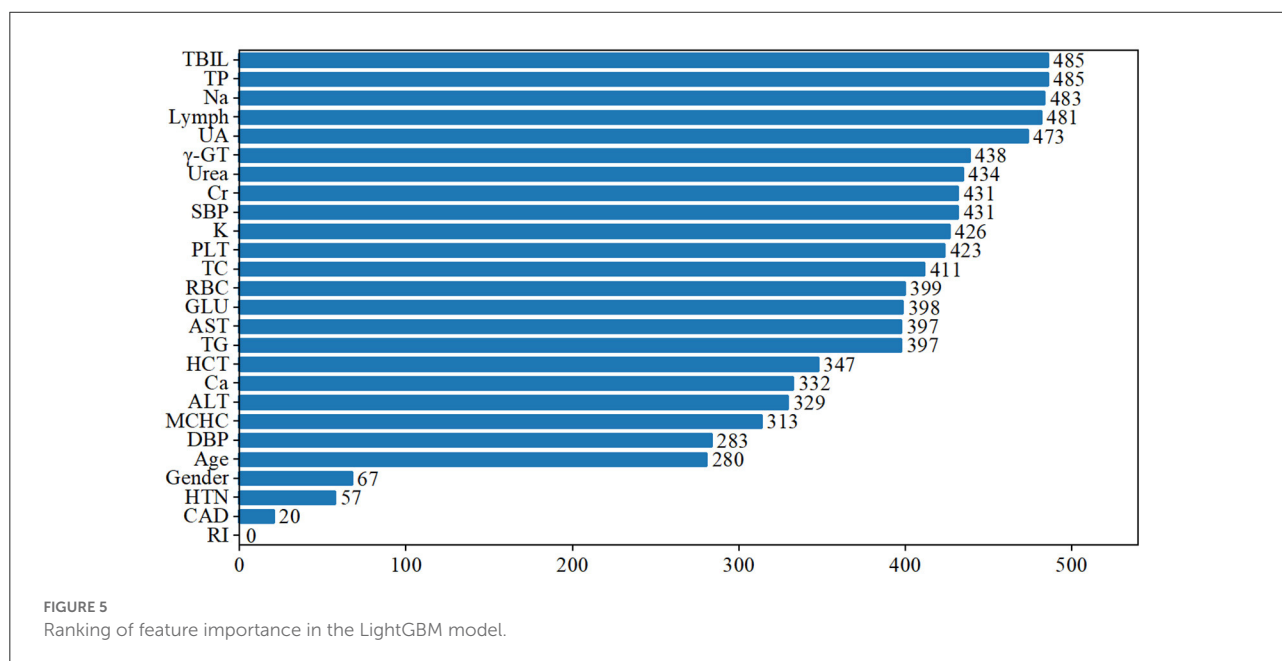
We compared five ML algorithms, and all were tested using a 5-fold cross-validation approach. In our study, the LightGBM model had a micro-F1 value of 78.4%, significantly higher than the other four ML models. These results demonstrate the superiority of the LightGBM model for determining which antihypertension drugs will be most effective in the elderly. LightGBM is an integrated model of decision tree for classification and regression prediction. It has an excellent version in disease diagnosis, such as Rufo (46) using LightGBM

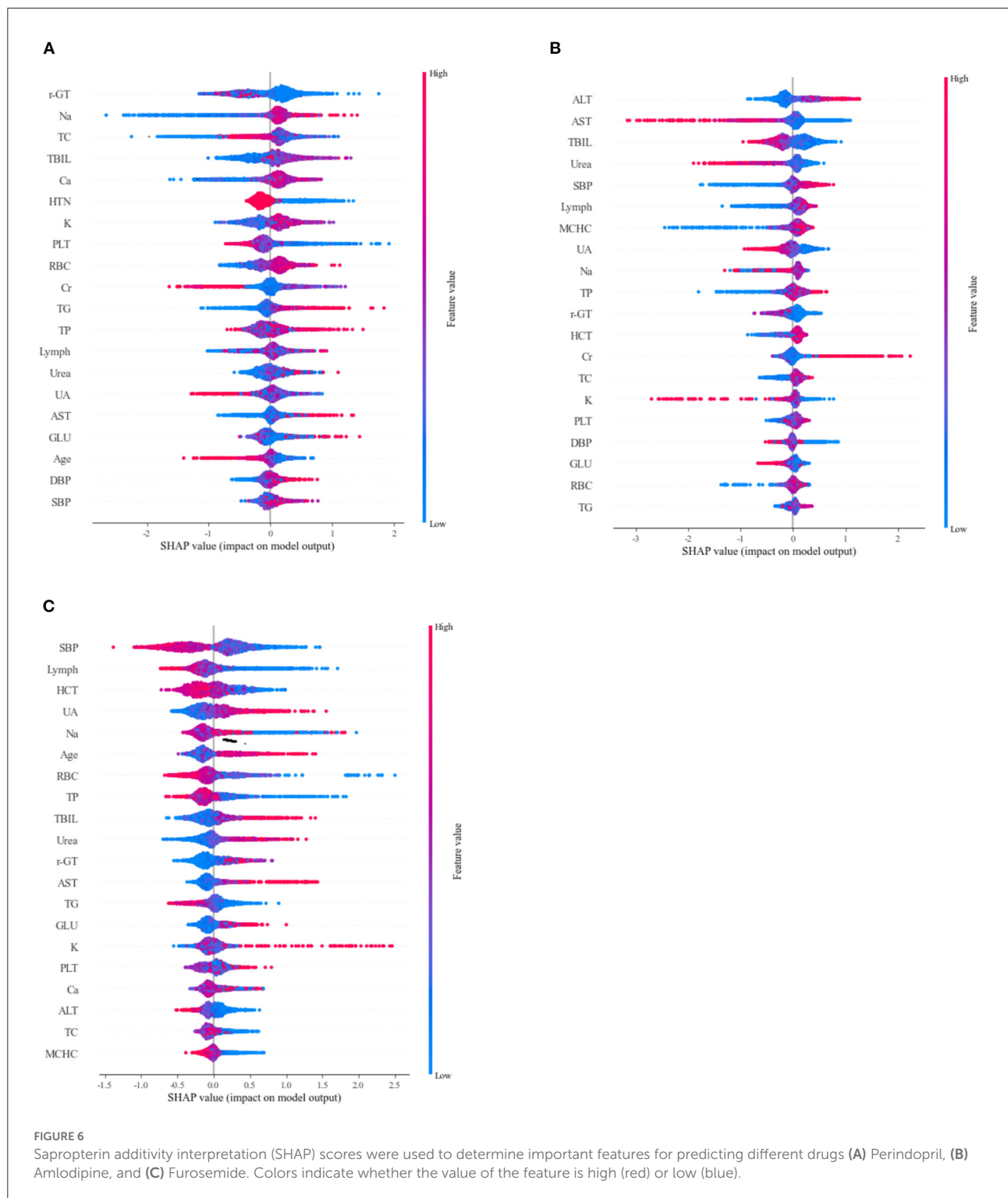
to predict diabetes, and Chen (47) proposed a weighted bagging-LightGBM model based on network embedding and PU learning for lncRNA disease association prediction. RF achieved great precision and micro-F1, 77.7%, behind only LightGBM. It is widely accepted that RF is the most useful integrated learning approach for both classification and regression (48). RF offers excellent classification accuracy, is tolerant to outliers and noise, and does not engage in overfitting (49). Both Subasi (50) and Singh (51) employed RF models to diagnose chronic renal disease and forecast cardiovascular disease, respectively. Besides, SVM, ANN, and NB are often applied in therapeutic situations. Shankar (52) developed an optimal multicore SVM model for feature-based thyroid illness classification. Radhimeenakshi (53) used SVM and ANN to classify coronary heart disease. The number of cases of coronavirus disease in China, Japan, Singapore, Iran, Italy, South Africa, and the United States in 2019 was predicted using an ANN model built by Nazar (54). Montazeri (55) created an NB-based prediction model to predict survival in various breast cancer types. To further explain the model, we also computed the SHAP values of the top-performing model, allowing us to examine the effect of the characteristics of each cause type on the model output.

Undoubtedly, the findings of this research have provided some groundwork for the following study. However, the current research does have a few limitations. This research is retrospective, and there is bias in the data, which is an issue that arises in all retrospective studies. On the other hand, the treatment of hypertension in clinical practice is mainly based on administering many medications in

TABLE 3 Comparison of five ML models.

| Model | Precision | Recall | F1-score |
|----------|-----------|--------|----------|
| RF | 0.777 | 0.777 | 0.777 |
| SVM | 0.731 | 0.731 | 0.731 |
| LightGBM | 0.784 | 0.784 | 0.784 |
| ANN | 0.724 | 0.724 | 0.724 |
| NB | 0.716 | 0.716 | 0.716 |





conjunction with one another. Only three kinds of single-drug modeling with greater frequency were chosen for the prediction investigation in this particular research, which is obviously

not enough. In the following sections, we will discuss the multi-label classification, which may accurately forecast various treatment regimens.

Conclusion

This study constructed a single-drug prediction model for antihypertensive drugs based on SVM, RF, LightGBM, ANN, and NB algorithms. And we also compared the efficacy of the model with evaluation metrics such as accuracy, recall, and micro-F1 to reduce the bias caused by a single algorithm and a single evaluation metric to some extent. The results showed that LightGBM predicted the best results compared with other algorithms, with a micro-F1 value of 78.4%. We also used feature importance ranking and SHAP values to enhance the interpretability of the model. We suggest that “increased Na, TBIL, and Ca, and decreased γ -GT and TC” can be prioritized for Perindopril, and “higher ALT and SBP, lower AST, TBIL, and Urea” can be prioritized with Amlodipine, “higher levels of UA, low levels of SBP, Lymph, HCT, and Na” can be prioritized for Furosemide. The study included patients from multiple medical institutions, and the study’s results may help clinicians provide decision support in prescribing antihypertensive drugs to patients.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

TW and WZ contributed to the conception and drafting of the study. YY helped to develop new techniques and programs.

References

- Oparil S, Acelajado MC, Bakris GL, Berlowitz DR, Cifková R, Dominiczak AF, et al. Hypertension. *Nat Rev Dis Primers*. (2018) 4:18014. doi: 10.1038/nrdp.2018.14
- B Banegas JR. [Epidemiology of hypertension in the elderly population of Spain]. *Anales de la Real Academia Nacional de Medicina*. (2002) 119:143–52.
- Peltzer K, Phaswana-Mafuya N. Hypertension and associated factors in older adults in South Africa. *Cardiovasc J Afr*. (2013) 24:66–72. doi: 10.5830/CVJA-2013-002
- Skliros EA, Papaioannou I, Sotiropoulos A, Giannakaki G, Milingou M, Lionis C. A high level of awareness but a poor control of hypertension among elderly Greeks. The Nemea primary care study. *J Hum Hypertens*. (2002) 16:285–7. doi: 10.1038/sj.jhh.1001380
- Triantafyllou A, Douma S, Petidis K, Doumas M, Panagopoulou E, Pырpasopoulou A, et al. Prevalence, awareness, treatment and control of hypertension in an elderly population in Greece. *Rural Remote Health*. (2010) 10:1225. doi: 10.22605/RRH1225
- Méndez-Chacón E, Santamaría-Ulloa C, Rosero-Bixby L. Factors associated with hypertension prevalence, unawareness and treatment among Costa Rican elderly. *BMC Public Health*. (2008) 8:275. doi: 10.1186/1471-2458-8-275
- NCD Risk Factor Collaboration (NCD-RisC). Long-term and recent trends in hypertension awareness, treatment, and control in 12 high-income countries: an analysis of 123 nationally representative surveys. *Lancet*. (2019) 394:639–51. doi: 10.1016/S0140-6736(19)31145-6
- Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL, et al. The seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure: the JNC 7 report. *JAMA*. (2003) 289:2560–72. doi: 10.1001/jama.289.19.2560
- Hajjar I, Kotchen TA. Trends in prevalence, awareness, treatment, and control of hypertension in the United States, 1988–2000. *JAMA*. (2003) 290:199–206. doi: 10.1001/jama.290.2.199
- NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants. *Lancet*. (2021) 398:957–80. doi: 10.1016/S0140-6736(21)01330-1
- Weisser B, Mengden T, Düsing R, Vetter H, Vetter W. Normal values of blood pressure self-measurement in view of the 1999 World Health Organization-International Society of Hypertension guidelines. *Am J Hypertens*. (2000) 13:940–3. doi: 10.1016/S0895-7061(00)00259-4

Data analysis by TW, SX, JT, and CY. The final text was reviewed and approved by all writers.

Funding

This study has been financially supported by the College of Medical Informatics, Chongqing Medical University, China, Student Research and Innovation Experiment Project (YJSZHYX202120).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcvm.2022.1056263/full#supplementary-material>

12. Tabesh M, Magliano DJ, Tanamas SK, Surmont F, Bahendeka S, Chiang C-E, et al. Cardiovascular disease management in people with diabetes outside North America and Western Europe in 2006 and 2015. *Diabet Med.* (2019) 36:878–87. doi: 10.1111/dme.13858
13. Ball R, Robb M, Anderson S, Dal Pan G. The FDA's sentinel initiative-A comprehensive approach to medical product surveillance. *Clin Pharmacol Ther.* (2016) 99:265–8. doi: 10.1002/cpt.320
14. Jernberg T, Attebring MF, Hambraeus K, Ivert T, James S, Jeppsson A, et al. The Swedish web-system for enhancement and development of evidence-based care in heart disease evaluated according to recommended therapies (SWEDEHEART). *Heart.* (2010) 96:1617–21. doi: 10.1136/hrt.2010.198804
15. Cowie MR, Blomster JJ, Curtis LH, Duclaux S, Ford I, Fritz F, et al. Electronic health records to facilitate clinical research. *Clin Res Cardiol.* (2017) 106:1–9. doi: 10.1007/s00392-016-1025-6
16. Pei Z, Liu J, Liu M, Zhou W, Yan P, Wen S, et al. Risk-predicting model for incident of essential hypertension based on environmental and genetic factors with support vector machine. *Interdiscip Sci Comput Life Sci.* (2018) 10:126–30. doi: 10.1007/s12539-017-0271-2
17. Golino HF, Souza D, Duarte SP, Gomes CA, Telma D, Reis LD, et al. Predicting increased blood pressure using machine learning. *J Obes.* (2014) 2014:637635. doi: 10.1155/2014/637635
18. Wang M-H, Li N-F, Luo Q, Wang G-L, Heizhati M, Wang L, et al. Development and validation of a novel diagnostic nomogram model to predict primary aldosteronism in patients with hypertension. *Endocrine.* (2021) 73:682–92. doi: 10.1007/s12020-021-02745-7
19. Gong J, Du C, Zhong X, Xiang T, Wang H. Researches on the illness risk of essential hypertension complicated with coronary heart disease based on machine learning algorithm. *Med J Chin Peoples Liberation Army.* (2020) 7:735–41. doi: 10.11855/j.issn.0577-7402.2020.07.10
20. Koren G, Nordon G, Radinsky K, Shalev V. Machine learning of big data in gaining insight into successful treatment of hypertension. *Pharmacol Res Perspect.* (2018) 6:e00396. doi: 10.1002/prp2.396
21. Chunyu L, Ran L, Junteng Z, Miye W, Jing X, Lan S, et al. Characterizing the critical features when personalizing antihypertensive drugs using spectrum analysis and machine learning methods. *Artif Intellig Med.* (2020) 104:101841. doi: 10.1016/j.artmed.2020.101841
22. Liu C, Xu J, Liu R, Wang M, Zhuo Y, Su L, et al. Identifying drug-related attributes to personalise antihypertensive agents: the outcome report of patients receiving metoprolol therapy. *BMC Med Inform Decis Mak.* (2021) 21:370. doi: 10.1186/s12911-021-01739-9
23. Guyon IM, Andr E. An introduction to variable and feature selection. *J Mach Learn Res.* (2003) 3:1157–82. doi: 10.1063/1.106515
24. Agodoa LY. Effect of ramipril vs amlodipine on renal outcomes in hypertensive nephrosclerosis: a randomized controlled trial. *JAMA.* (2001) 285:2719. doi: 10.1001/jama.285.21.2719
25. Jennings GLR. Recent clinical trials of hypertension management. *Hypertension.* (2013) 62:3–7. doi: 10.1161/HYPERTENSIONAHA.113.00863
26. Williams B, Mancia G, Spiering W, Agabiti Rosei E, Azizi M, Burnier M, et al. 2018 ESC/ESH Guidelines for the management of arterial hypertension: the Task Force for the management of arterial hypertension of the European Society of Cardiology (ESC) and the European Society of Hypertension (ESH). *Eur Heart J.* (2018) 39:3021–104. doi: 10.1097/HJH.0000000000001961
27. Whelton PK, Carey RM, Aronow WS, Casey DE, Wright JT. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults. A report of the American College of Cardiology/American Heart Association Task Force on Clinical Pra. *Hypertension.* (2017) 19:213–21. doi: 10.1161/HYP.0000000000000065
28. Julius S, Kjeldsen SE, Weber M, Brunner HR, Ekman S, Hansson L, et al. Outcomes in hypertensive patients at high cardiovascular risk treated with regimens based on valsartan or amlodipine: the VALUE randomised trial. *Lancet.* (2004) 363:2022–31. doi: 10.1016/S0140-6736(04)16451-9
29. The ALLHAT Officers and Coordinators for the ALLHAT Collaborative Research Group. Major outcomes in high-risk hypertensive patients randomized to angiotensin-converting enzyme inhibitor or calcium channel blocker vs diuretic: the antihypertensive and lipid-lowering treatment to prevent heart attack trial (ALLHAT). *JAMA.* (2002) 288:2981–97. doi: 10.1001/jama.288.23.2981
30. Dahlöf B, Sever PS, Poulter NR, Wedel H, Beevers DG, Caulfield M, et al. Prevention of cardiovascular events with an antihypertensive regimen of amlodipine adding perindopril as required versus atenolol adding bendroflumethiazide as required, in the Anglo-Scandinavian Cardiac Outcomes Trial-Blood Pressure Lowering Arm (ASCOT-BPLA): a multicentre randomised controlled trial. *Lancet.* (2005) 366:895–906. doi: 10.1016/S0140-6736(05)67185-1
31. Vlase HL, Panagopoulos G, Michelis MF. Effectiveness of furosemide in uncontrolled hypertension in the elderly: role of renin profiling. *Am J Hypertens.* (2003) 16:187–93. doi: 10.1016/S0895-7061(02)03260-0
32. Wu Y, Wang X, Zhang Q. Protective effect of perindopril on target organ in elderly patients with mild-to-moderate hypertension. *Intern Med China.* (2015) 10:597–600. doi: 10.6121/j.cnki.cn45-1347/r.2015.05.01
33. Stolarz-Skrzypek K, Kuznetsova T, Thijs L, Tikhonoff V, Seidlerová J, Richart T, et al. Fatal and nonfatal outcomes, incidence of hypertension, and blood pressure changes in relation to urinary sodium excretion. *JAMA.* (2011) 305:1777–85. doi: 10.1001/jama.2011.574
34. Yang Y. Analysis of blood routine and some biochemical indexes in elderly hypertensive patients. *Int J Lab Med.* (2016) 37:3075–8. doi: 10.3969/j.issn.1673-4130.2016.21.054
35. Wang M. Influence of white blood cell count on disease control in elderly patients with essential hypertension. *J Med Forum.* (2015) 36:109–110. Available online at: https://xueshu.baidu.com/usercenter/paper/show?paperid=8c7e31d95dcca0a194ca0d6c3320c&site=xueshu_se&hitarticle=1
36. Gagnon DR, Zhang T-J, Brand FN, Kannel WB. Hematocrit and the risk of cardiovascular disease—the Framingham Study: a 34-year follow-up. *Am Heart J.* (1994) 127:674–82. doi: 10.1016/0002-8703(94)90679-3
37. Wei Y, Wang H, Hu Y, Gen Y, Song Q. Serum total protein and its related factorin very elderly hypertensive patients. *Chin J Integr Med Cardio Cerebrovasc Dis.* (2014) 12:1045–7. doi: 10.3969/j.issn.1672-1349.2014.09.007
38. Wang X, He J, Fan Q. The relationship between level of serum uric acid, total bilirubin and hypertension. *Clin Ration Drug Use.* (2012) 5:8. doi: 10.15887/j.cnki.13-1389/r.2012.09.059
39. Kunutsor SK, Kieneker LM, Burgess S, Bakker SJL, Dullaart RPF. Circulating total bilirubin and future risk of hypertension in the general population: the prevention of renal and vascular end-stage disease (PREVEND) prospective study and a Mendelian randomization approach. *J Am Heart Assoc.* (2017) 6:e006503. doi: 10.1161/JAHA.117.006503
40. Intersalt: an international study of electrolyte excretion and blood pressure. Results for 24 hour urinary sodium and potassium excretion. Intersalt Cooperative Research Group. *BMJ.* (1988) 297:319–28. doi: 10.1136/bmj.297.6644.319
41. Alderman MH, Cohen H, Madhavan S. Sodium reduction and weight loss for elderly patients with hypertension. *JAMA.* (1998) 280:885–6. doi: 10.1001/jama.280.10.885
42. Huang L, Trieu K, Yoshimura S, Neal B, Woodward M, Campbell NRC, et al. Effect of dose and duration of reduction in dietary sodium on blood pressure levels: systematic review and meta-analysis of randomised trials. *BMJ.* (2020) 368:m315. doi: 10.1136/bmj.m315
43. Stamler J, Chan Q, Daviglus ML, Dyer AR, Van Horn L, Garside DB, et al. Relation of dietary sodium (salt) to blood pressure and its possible modulation by other dietary factors: the INTERMAP study. *Hypertension.* (2018) 71:631–7. doi: 10.1161/HYPERTENSIONAHA.117.09928
44. Guzik TJ, Hoch NE, Brown KA, McCann LA, Rahman A, Dikalov S, et al. Role of the T cell in the genesis of angiotensin II-induced hypertension and vascular dysfunction. *J Exp Med.* (2007) 204:2449–60. doi: 10.1084/jem.20070657
45. Gao H, Li L. Research progress in the relationship between neutrophil/lymphocyte and essential hypertension. *Hainan Med J.* (2018) 29:2788–90. doi: 10.3969/j.issn.1003-6350.2018.19.037
46. Rufo DD, Debelee TG, Ibenthal A, Negera WG. Diagnosis of diabetes mellitus using gradient boosting machine (LightGBM). *Diagnostics.* (2021) 11:1714. doi: 10.3390/diagnostics11091714
47. Chen X, Liu X. A Weighted Bagging LightGBM Model for Potential lncRNA-Disease Association Identification. In: Qiao J, Zhao X, Pan L, Zuo X, Zhang X, Zhang Q, et al., editors. *Bio-inspired Computing: Theories and Applications. Communications in Computer and Information Science.* Singapore: Springer Singapore (2018). p. 307–14. doi: 10.1007/978-981-13-2826-8_27
48. Wang Y, Xia S-T, Tang Q, Wu J, Zhu X. A novel consistent random forest framework: Bernoulli random forests. *IEEE Trans Neural Netw Learn Syst.* (2018) 29:3510–23. doi: 10.1109/TNNLS.2017.2729778
49. Liu Y, Wang Y, Zhang J. New machine learning algorithm: random forest. In: Liu B, Ma M, Chang J, editors. *Information Computing and Applications. Lecture Notes in Computer Science.* Berlin, Heidelberg: Springer (2012). p. 246–52. doi: 10.1007/978-3-642-34062-8_32
50. Subasi A, Alickovic E, Kevric J. Diagnosis of chronic kidney disease by using random forest. In: Badnjevic A, editor. *CMBEBIH 2017. IFMBE Proceedings.* Singapore: Springer (2017). p. 589–94. doi: 10.1007/978-981-10-4166-2_89
51. Singh YK, Sinha N, Singh SK. Heart disease prediction system using random forest. In: Singh M, Gupta PK, Tyagi V, Sharma A, Ören T,

Grosky W, editors. *Advances in Computing and Data Sciences. Communications in Computer and Information Science*. Singapore: Springer (2017). p. 613–23. doi: 10.1007/978-981-10-5427-3_63

52. Shankar K, Lakshmanaprabu SK, Gupta D, Maselena A, de Albuquerque VHC. Optimal feature-based multi-kernel SVM approach for thyroid disease classification. *J Supercomput.* (2020) 76:1128–43. doi: 10.1007/s11227-018-2469-4

53. Radhimeenakshi S. Classification and prediction of heart disease risk using data mining techniques of support vector machine and artificial neural network.

In: *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*. New Delhi: IEEE (2016). p. 3107–111.

54. Niazkar HR, Niazkar M. Application of artificial neural networks to predict the COVID-19 outbreak. *Glob Health Res Policy.* (2020) 5:50. doi: 10.1186/s41256-020-00175-y

55. Montazeri M, Montazeri M, Montazeri M, Beigzadeh A. Machine learning models in breast cancer survival prediction. *Technol Health Care.* (2016) 24:31–42. doi: 10.3233/THC-151071