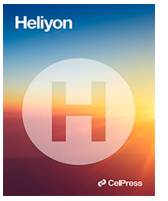
Heliyon 10 (2024) e37635

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Enhancing diabetic foot ulcer prediction with machine learning: A focus on Localized examinations☆

Wang Xiaoling [a,1], Zhu Shengmei [b,1], Wang BingQian [c,1], Li Wen [a], Gu Shuyan [d], Chen Hanbei [a], Qin Chenjie [e,**], Dai Yao [f,***], Li Jutang [g,*]

[a] *Department of Endocrinology, Xinhua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai 200092, China*
[b] *Department of Pediatric Surgery, Xin Hua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai 200092, China*
[c] *Intensive Care Medicine Department, Suzhou Traditional Chinese Medicine Hospital, Suzhou, Jiangsu 215009, China*
[d] *Center for Health Policy and Management Studies, School of Government, Nanjing University, Nanjing 210023, China*
[e] *Department of Oncology, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, 200080, China*
[f] *Nursing Department of Xinhua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai 200092, China*
[g] *Hongqiao International Institute of Medicine,Tongren Hospital, Shanghai Jiao Tong University School of Medicine, 1111 XianXia Road, Shanghai 200336, China*

A R T I C L E I N F O

Keywords:
Diabetic foot ulcer (DFU)
Machine learning
Predict model
Local foot examinations
Screening

A B S T R A C T

Background: diabetices foot ulcer (DFU) are serious complications. It is crucial to detect and diagnose DFU early in order to provide timely treatment, improve patient quality of life, and avoid the social and economic consequences. Machine learning techniques can help identify risk factors associated with DFU development.

Objective: The aim of this study was to establish correlations between clinical and biochemical risk factors of DFU through local foot examinations based on the construction of predictive models using automated machine learning techniques.

Methods: The input dataset consisted of 566 diabetes cases and 50 DFU risk factors, including 9 local foot examinations. 340 patients with Class 0 labeling (low-risk DFU), 226 patients with Class 1 labeling (high-risk DFU). To divide the training group (consisting of 453 cases) and the validation group (consisting of 113 cases), as well as preprocess the data and develop a prediction model, a Monte Carlo cross-validation approach was employed. Furthermore, potential high-risk factors were analyzed using various algorithms, including Bayesian BYS, Multi-Gaussian Weighted Classifier (MGWC), Support Vector Machine (SVM), and Random Forest Classifier (RF). A three-layer machine learning training was constructed, and model performance was estimated using a Confusion Matrix. The top 30 ranking feature variables were ultimately determined. To reinforce the robustness and generalizability of the predictive model, an independent dataset comprising 248 cases was employed for external validation. This validation process evaluated the model's applicability and reliability across diverse populations and clinical

☆ This study was supported by the Shanghai Municipal Education Commission-Gaoyuan Nursing Grant Support (Hlgy1810kyx), Xin Hua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine Nursing Development Program (HLXKRC2021004), Innovative Research Team of High-level Local Universities in Shanghai (SHSMU-ZLCX20212401) and National Natural Science Foundation of China (72104102).

* Corresponding author.
** Corresponding author.
*** Corresponding author.
E-mail addresses: 13661919158@163.com (Q. Chenjie), 1374760631@qq.com (D. Yao), jutangli@shsmu.edu.cn (L. Jutang).

[1] Both authors contributed equally to this work.

https://doi.org/10.1016/j.heliyon.2024.e37635
Received 29 April 2024; Received in revised form 3 September 2024; Accepted 6 September 2024
Available online 19 September 2024

settings. Importantly, the external dataset required no additional tuning or adjustment of parameters, enabling an unbiased assessment of the model's generalizability and its capacity to predict the risk of DFU.

Results: The ensemble learning method outperformed individual classifiers in various performance evaluation metrics. Based on the ROC analysis, the AUC of the AutoML model for assessing diabetic foot risk was 88.48 % (74.44–97.83 %). Other results were found to be as follows: 87.23 % (63.33 %–100.00 %) for sensitivity, 87.43 % (70.00 %–100.00 %) for specificity, 87.33 % (76.66 %–95.00 %) for accuracy, 87.69 % (75.00 %–100.00 %) for positive predictive value, and 87.70 % (71.79 %–100.00 %) for negative predictive value. In addition to traditional DFU risk factors such as cardiovascular disorders, peripheral artery disease, and neurological damage, we identified new risk factors such as lower limb varicose veins, history of cerebral infarction, blood urea nitrogen, GFR (Glomerular Filtration Rate), and type of diabetes that may be related to the development of DFU. In the external validation set of 158 samples, originating from an initial 248 with exclusions due to missing labels or features, the model still exhibited strong predictive accuracy. The AUC score of 0.762 indicated a strong discriminatory capability of the model. Furthermore, the Sensitivity and Specificity values provided insights into the model's ability to correctly identify both DFU cases and non-cases, respectively.

Conclusion: The predictive model, developed through AutoML and grounded in local foot examinations, has proven to be a robust and practical instrument for the screening, prediction, and diagnosis of DFU risk. This model not only aids medical practitioners in the identification of potential DFU cases but also plays a pivotal role in mitigating the progression towards adverse outcomes. And the recent successful external validation of our DFU risk prediction model marks a crucial advancement, indicating its readiness for clinical application. This validation reinforces the model's efficacy as an accessible and reliable tool for early DFU risk assessment, thereby facilitating prompt intervention strategies and enhancing overall patient outcomes.

1. Introduction

DFU represent a significant complication of diabetes mellitus (DM) [1], characterized by a high incidence and recurrence rate. The International Diabetes Foundation estimates that 40 million to 60 million people globally are affected by DFU [2], a marked increase from 2015 estimates that ranged from 9 million to 26 million [3], Furthermore, the recurrence rate of DFU within one year remains high, at approximately 40 %, despite successful treatment [4]. In China, the incidence of DFU within 1 year in DM patients is 8.1 % [5], and the recurrence rate of healed DFU within 12 weeks is also as high as 60 % [6]. Other characteristics of DFU include high disability and mortality. DM patients are most likely to suffer from severe DFU that cannot be healed, and it is also primary cause of non-traumatic limb amputations [7]. In China, individuals with DFU who are over 50 years of age have an incidence of amputation as high as 22 % [5]. On the other hand, a cohort study found that DFU was independently associated with 5 % mortality within 1 year and 42 % mortality within 5 years in DM patients. Patients with DFU had a 2–3 times higher death rate than those without DFU [8]. In addition, the five-year mortality rate among individuals with DFU who undergo amputation is as high as 80 %, far higher than that of most malignancies [9]. DM medical costs in China are expected to increase from $4.9 billion to $7.4 billion by 2030, with DFU accounting for 20 % of the total [10]. Therefore, it is particularly important to effectively manage DM patients and reduce the incidence of DFU in a country with a large population like China.

The early detection and timely intervention of DFU is one of the difficulties in DM management. Long-term persistent hyperglycemic state can impair neurological functions, and autonomic neuropathy can lead to reduced sweat secretion, resulting in dry and chapped skin and deep skin structures susceptible to microbial invasion and colonization [11]. In addition, in patients with type 2 DM, disorders of lipid metabolism could promote atherosclerosis, leading to arterial narrowing and reduced blood supply to the foot [12]. The above pathological processes diminish foot sensation, and the unobserved trauma will accelerate DFU formation. Besides, immune dysregulation and persistent inflammatory stress [13,14] lead to decreased immune defense of the body which makes slow wound healing and recurrent infections, finally resulting in amputation, and even death [15]. Therefore, the early detection and management of DFU is a pressing issue at present.

DFU is currently managed through patient education, self-management, regular foot examinations, and comprehensive medical assessment based on the medical history of the patient and the peripheral vascular and sensory functions of the patient. In spite of this, the main risk factors for DFU are not well understood [6]. The self-management could be limited due to diminished sensation or inadequate knowledge reserve [16]. In addition, the assessment methods for DFU are not uniform and time-consuming, and the extent of DFU vascular lesions, neuropathy, infection, and soft tissue and bone destruction substantially varies. Thus, the assessment mainly relies on clinical experience, which is more subjective than objective statistical methods, so its extension may be also limited [17]. Besides, the statistical methods traditionally used for risk assessment in patients with DFU are mainly linear models with simple algorithms that may be unable to analyze the complex nonlinear relationships between risk factors [18]. In addition to age, gender [2,19,20], smoking, BMI, foot pulse, and Semmes-Weinstein monofilament examination, there are more than 20 risk factors involved in DFU [21], and the neglect of their inner relationships may lead to impaired predictive ability and inaccurate predictive values [22]. The classification systems provide clinicians with a wider differential diagnosis for DFU by including both neuropathy and ischemia [23,24].

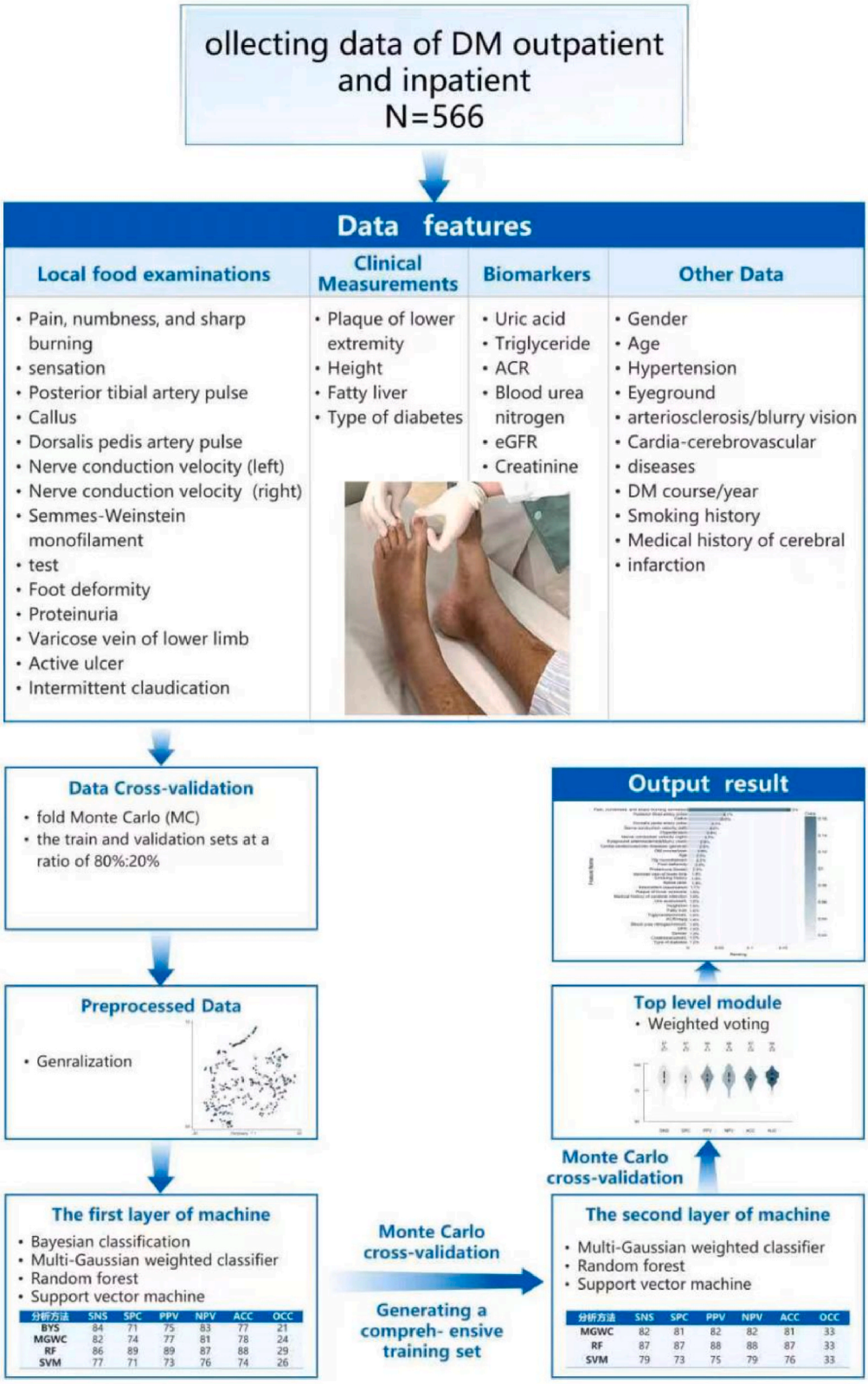


Fig. 1. The methodology of this study.

Timely diagnosis is essential.A multicentered study across Europe demonstrated that delay is common and thus a possible rate limiting factor in providing effective treatment for these patients as delays allow the progression of foot disease and a possible lost opportunity [25]. Based on the uneven distribution of medical resources, large differences in medical levels, the shortage of medical personnel in China [26], and the aforementioned reasons, it is particularly important to construct a DFU risk warning system and a management model for chronic diseases which are relatively simple to operate and easy to popularize.

Automated machine learning (AutoML), a branch of artificial intelligence, can be used for the diagnosis, prognostic observation and management of diseases with the advantages like less human errors, cost, and higher adaptability compared to traditional linear models. More and more compelling evidence suggests that AutoML holds great promise for improving the accuracy and reliability of clinical decision making [27]. Studies have shown the integration of clinical diagnosis-supporting system with the self-management of DM patients could be achieved by wearable smart devices [28] to realize carbohydrate intake calculation, blood glucose monitoring and prediction, drug compliance, physical activity monitoring, early risk event prediction (including hyperglycemia and hypoglycemia), and automatic insulin dose adjustment [29,30]. Sneha et al. [31] based on 11 features constructed a predictive model by Naive Bayes Algorithm, applied it on a DM sample dataset (n = 2500), and the results showed the highest accuracy of 82.30 % for the test results.

In a study by Arcadu et al. [32] deep learning was applied to predict diabetic retinopathy progression, and maximum sensitivity (79 ± 12 %) was reached at the 12th month while maximum specificity (77 ± 12 %) was reached at the 6th month. These show the significant advantages of AutoML in prediction, smart medical care and self-management of DM management. Utilizing machine learning techniques to process the thermographic features of diabetic foot patients also contributes to the high-precision detection and identification of diabetic foot [33,34]. Although there have been previous AutoML methods combined with computer vision for DFU high-risk identification [35,36], there are no data of DFU high-risk people and reports on the construction of DFU risk prediction models. Meanwhile, the above studies also provide a good preliminary basis and theoretical basis for our in-depth research on AutoML in the prediction and management of DFU high-risk people.

In this study, based on an existing dataset of diabetic foot patients, we employed a novel technique comprised of multiple integrated algorithms to evaluate features obtained to enhance the diagnosis of diabetic foot. The main contributions of this study include.

- Integration of local diabetic foot examination with clinical and biochemical feature characteristics to achieve high accuracy detection of diabetic foot by combining local and global features.
- Thorough investigation of relevant features to improve detection performance, where these features are utilized as inputs in traditional classifiers.
- Exploration of feature selection and optimization techniques, as well as classification models, to maximize detection performance using lightweight classifiers.

2. Methods

The subjects of this study come from the Department of Endocrinology of our hospital, including patients who were previously hospitalized or treated as outpatients, diagnosed with diabetes, with or without complications of diabetic foot ulcer. All patients in this study retrospectively included from February 2020 to April 2021. 566 patients were included, including 308 males and 258 females, age 65 (59, 71), diabetic duration 11 (4, 17). This study involved voluntary participation and informed consent from all participants. Patients with (i) severe liver or kidney disease or other critical illnesses (e.g. malignancy), (ii) incomplete or severely missing diagnostic information, or (iii) refusal to participate in this study were excluded. Clinical information was collected from participants, including 50 risk indicators such as peripheral nerve examination, peripheral artery examination, foot deformity, the history of DFU and kidney failure. We have identified multiple risk factors contributing to the occurrence of diabetic foot ulcers, including diabetic complications, comorbidities, abnormal laboratory parameters, and demographic data. Through literature validation, we have identified foot local examination as one of the primary risk factors. Participants were also scored using the DFU risk rating scale [37], with a score of 0 for low risk people and 1 for high risk people. Meantime, the reference categories used for cross-validation were category 0 (low-risk DFU) or category 1 (high-risk DFU) subgroups [38]. No biospecimens were collected from patients in this study. The confidentiality of patients' privacy and identity information was guaranteed, and there was no commercial interest involved, so that this research met the conditions for exemption from informed consent. Our hospital's ethics committee approved this research (reference:XHEC-D-2024-069, date: 2024-4-25). Fig. 1 for details of the methodology.

A flowchart showing how machine learning models are developed and deployed.The framework of predictive model based on local foot examinations. This study was to establish correlations between clinical and biochemical risk factors of DFU through local foot examinations.

2.1. Cross-validation

Cross-validation using 100-fold Monte Carlo (MC) [39], patients in each chort were randomly assigned to training and validation sets, with a ratio of 80 % training to 20 % validation. This process was repeated 100 times to minimize the variance in predictive performance estimation. Consequently, each chort yielded 100 unique training-validation configurations. MC split resulted in 453 samples per fold in the training set and 113 samples per fold in the validation set. Subsequently, an equal subsampling was performed in the validation set to ensure no overestimation or underestimation of reference categories during the cross-validation process.

2.2. Data pre-processing

Since the dataset has missing values or outliers, we take pre-processing of the data before performing the prediction training on the training set, which is able to handle inconsistent data and missing values.

2.2.1 The initial stage involves data normalization, which refers to the process of adjusting the range of the original data to a specific set of values. The purpose is to eliminate the dimensional differences between different features, and this process is achieved through Feature Normalization (FN). In this dataset, we apply variance reduction to improve the accuracy of classifier [40,41].

2.2.2 In the second step, kernel functions are used to solve nonlinear problems. Kernel functions are the bridge between linear and nonlinear problems, and can be utilized to calculate the similarity between samples. It encompasses various function types such as polynomial, sigmoid, radial basis function (RBF), linear and so on [42]. In this study, polynomial kernel function and RBF are chosen for the following reasons, (i) polynomials are able to handle image data and suitable for problems of totally normalized training data; (ii) under the condition of no increasing complexity, RBF and polynomial kernel function are more suitable for nonlinear problems like the prediction of DFU. Meanwhile, RBF is applicable for problems without a priori knowledge of the relevant data [43,44].

2.2.3 In the third step, considering that many characteristic variables may appear weakly correlated with the occurrence of DFU, we don't apply all of them in the prediction model [45]. In sequential forward selection methods, characteristic variables can belong to multiple categories, and representative features are chosen from each orthogonal category, which can span a more diverse feature space and help to exclude redundant information [46]. In this study, we adopted Pearson correlation analysis before selecting the characteristic variables, and then correlation matrix analysis to remove redundant features, this process can be achieved through Smart Redundancy Reduction (SRR) algorithm, Ensuring the similarity of sample features while eliminating redundancy will further enhance the model's generalization and interpretability. Features with a Pearson correlation coefficient exceeding 0.85 were deemed redundant. And the data after removing redundant features were subjected to the machine learning modeling [47]. Subsequently, the average of all the characteristic variables ranking after MC validation was calculated, and then the final importance of the characteristic variables was determined based on the calculated results. The top 30 characteristic variables were selected. The preprocessing steps and their parameters are detailed in Table 1.

2.3. The construction of predictive algorithm

In order to minimize the impact of method bias and improve prediction accuracy, we apply ensemble learning with four algorithms to construct the predictive model, including Bayesian (BYS), Multi-Gaussian Weighted Classifier (MGWC), Support Vector Machine (SVM) algorithm, and Random Forest Classifier (RF).

2.3.1. Classifier

2.3.1.1 BYS is a simple, efficient and supervised learning method for classifying unlabeled data based on probabilities [48]. BYS assumes that the distribution of the model variables follows some probability, and arrives at the optimal decision from the probability distribution and the observed data. It has a solid mathematical foundation and a relatively stable classification capability because of its core Bayes' theorem. BYS requires no iteration in the training phase with a relatively low time complexity, and it assumes that the attribute conditions are independent of any other features. Moreover, BYS has a high computational speed in the modeling process, which can greatly simplify learning [49].

2.3.1.2 MGWC is used in machine learning to classify observations into multivariate normal classes and to model Gaussian processes using variances. Specifically, each category is modelled as a Gaussian distribution, whose parameters are used to describe its feature distribution, such as mean and covariance matrices. As new data is being classified, the classifier calculates the distance between each data point and the Gaussian probability function of the category, and weighted it to determine which category the data points belong to Refs. [50–52].

2.3.1.3 SVM is a supervised learning-based generalized linear classifier that computes a hyperplane with maximum margin as the decision boundary. It maps data information in the form of points into a high-dimensional space and expands the spacing by the hyperplane with minimal empirical classification error as much as possible to differentiate classes [53,54]. The study's data samples comprise 30 features, each of which is a vector of a specific dimension. Subsequently, SVM is employed to construct a hyperplane in the hyperspace, enabling improved differentiation of the classes.

Table 1
Performing preprocessing steps, algorithms, parameters, and values before machine learning for all Monte Carlo folds.

Preprocessing step	Algorithm	Parameter	Value
1	FN	Normalization type	Mean-Deviation
2	KE	Kernels applied	Gaussian; Polynomial; Tanh
3	SRR	Redundancy Threshold (Covariance)	0.85

Preprocessing step algorithms as well as their parameter values performed in all Monte Carlo folds before machine learning. FN - Feature Normalization; KE - Kernel-based Feature Engineering; SRR - Smart Redundancy Reduction.

2.3.1.4 RF is a predictive algorithm based on nonparametric classification and regression for building decision tree ensemble [55]. Decision tree is a method of classifying data by randomly splitting the data at training phase based on the values of specific variables, and then repeating this split such that the split data set consists of target variables in the same class. RF generates numerous decision trees by randomly selecting subsets of a given dataset. Each tree is assigned a classification, and the classification with the most votes is used as the prediction model [55]. Two approaches can be employed to complete the model's prediction: if the output is the mean value, RF addresses the regression problem, while if it is a class pattern, RF tackles the classification problem [56,57]. Ultimately, the bagging method is used to combine multiple decision tree learning models, which are grown in a randomly selected subspace. Additionally, decision trees can determine the contribution of each variable to data classification and calculate its importance [58].

2.3.2. The steps of ensemble learning

2.3.2.1. The first layer of machine learning training. Initially, a range of machine learning algorithms were developed to mitigate the impact of algorithmic bias for each MC validation [47]. Subsequently, each model commenced training using pre-processed training data that was randomly selected. (Table 2 for detailed parameters of each algorithmic model).

2.3.2.2. The second layer of machine learning training. The Meta-training dataset (MTD) is created by evaluating the training models formed in the first ML layer on the samples from the pre-processed training set in each MC validation. To create it, The output of each training model in the initial layer of machine learning is employed as feature values for a given training sample. MTD trains the input of the prediction models in the second ML layer. The subsequent models are trained to identify the prediction patterns of the initial layer machine learning models, thereby producing hybrid super-learners [59]. The weights of each model in the second layer of machine learning are determined based on their training performance. Furthermore, models with training performance below the median of all second layer models are assigned a weight of 0 in the final vote (Table 3 for detailed parameters for each algorithmic model).

2.3.2.3. The third layer of machine learning training. The final predictive results of the model are obtained by applying weighted majority voting to the predictions generated by the second layer machine learning model. The importance of every feature in the prediction model is then estimated by calculating R-squared ranking in each cross-validation [60].

2.4. Performance evaluation

To assess model performance, including accuracy (ACC), sensitivity (SEN), specificity (SPE), positive predictive value (PPV), negative predictive value (NPV), receiver operating characteristic (ROC) curve, and area under curve (AUC), Confusion Matrix (CMM) analysis was employed. To obtain robust and balanced results, we performed cross-validation optimization three times on the machine learning, and each cross-validation optimization contained resampling to effectively train the prediction algorithm [61]. The decisions of final model were made based on the weighted votes of the three model predictions. Once constructed and trained, the final model will be independently validated by the validation set. During evaluation, the subsampling is also performed on the validation samples to ensure that no subgroup results are over- or underestimated during the cross-validation.

Table 2
The first layer of ML algorithms, along with their respective parameters and value ranges, were applied across Monte Carlo folds.

ML Algorithm	Parameter	Value Range	Occurre-nce
BYS	–	–	21.31 %
MGWC	Initial value multiplier	1–10	24.28 %
	Maximum iterations	29000–150000	
	Negative weights allowed	true, false	
	Scale value multiplier	0.1–50	
	Tolerance	0.00001–0.0001	
RF	Bag fraction	0.8–0.99	28.77 %
	Bagging method	equalized, normal	
	Boosting	adaboost, none	
	Maximum tree depth	5–23	
	Minimum samples in leaves	4–8	
	Node feature selection method	none	
	Number of random features per node	5	
	Number of selected trees	101–201	
	Number of trees to build	301–1001	
	Tree quality metric	gain, gini	
	Tree selection method	0	
SVM	Learning rate	0.001–0.01	25.62 %
	Maximum iterations	1000–5000	

Occurrence of each ML type is represented in percentages across MC folds. BYS: Bayesian Classifier; MGWC: Multi-Gaussian Weighted Classifier; RF: Random Forest Classifier; SVM: Support Vector Machine Classifier.

Table 3
The Second layer of ML algorithms with their corresponding parameters and value ranges through Monte Carlo folds.

ML Algorithm	Parameter	Value Range	Occurre-nce
MGWC	Initial value multiplier	10	33.33 %
	Maximum iterations	5000–17000	
	Negative weights allowed	true, false	
	Scale value multiplier	1–5	
	Tolerance	0.0001	
RF	Bag fraction	0.8–0.99	33.33 %
	Bagging method	equalized, normal	
	Boosting	adaboost, none	
	Maximum tree depth	5	
	Minimum samples in leaves	4–8	
	Node feature selection method	none	
	Number of random features per node	5	
	Number of selected trees	201	
	Number of trees to build	501–1001	
	Tree quality metric	gini, gain	
	Tree selection method	0	
SVM	Learning rate	0.001–0.01	33.33 %
	Maximum iterations	1000–5000	

Occurrence of each ML type is represented in percentages across MC folds. MGWC: Multi-Gaussian Weighted Classifier; RF: Random Forest Classifier; SVM: Support Vector Machine Classifier.

2.5. External validation of the predictive model

To reinforce the methodological robustness of our predictive model and to establish its applicability and reliability across different populations and clinical settings, we conducted an external validation using an independent dataset comprising 248 cases, including 147 males and 101 females, age 70 (66, 74), diabetic duration 10 (3, 18). excluding instances without labels and those with significant missing feature data. Missing values for features with less than three missing entries were imputed using the median value. The inclusion period for the external validation dataset is from December 2023 to February 2024. The external dataset required no further tuning or adjustment of parameters, which allowed us to assess the model's generalizability and its ability to predict DFU risk in an unbiased manner.

2.5.1. Data Preparation

Prior to validation, the dataset underwent rigorous preprocessing to align with the standards set by our initial model development. Data cleaning steps included handling of missing values, standardization of measurements, and normalization of features to ensure consistency with the training dataset.

2.5.2. Performance metrics

The external validation of the predictive model's performance was evaluated using the area under the receiver operating characteristic curve (AUC), sensitivity (SNS), specificity (SPC), positive predictive value (PPV), negative predictive value (NPV), and accuracy (ACC).

2.6. Statistical analysis

Continuous variables are presented as either mean ± standard deviation (SD) or median, while categorical variables are presented as numbers and percentages (%). Continuous variables were analyzed using the Mann-Whitney *U* test, while categorical variables were assessed using the chi-square test. Cor linear regression was conducted to explore the potential linear correlation between the indicators in the training set. To achieve two-dimensional visualization of high-dimensional data, we utilized t-distributed stochastic neighbor embedding (t-SNE), a nonlinear dimensionality reduction algorithm in machine learning, to visualize the results of the indicators obtained from routine laboratory tests. A p-value of less than 0.05 was considered statistically significant.

3. Results

3.1. Characteristics of the study objects

A total of 566 DM patients (4 cases of type 1 and 562 cases of type 2) from 15 to 95 years old (median age: 65 years old), including 308 males and 258 females, were recruited in this study. The patients with DM, whose duration of illness ranged from 1 to 42 years with a median duration of 11 years, were divided into the training group (453 cases) and validation group (113 cases) using the Monte Carlo (MC) cross-validation scheme. We pre-processed the data, constructed the prediction model, and finally presented the top 30 weighted variables. There were 340 patients labeled category 0 (low-risk DFU) and 226 patients labeled category 1 label (high-risk DFU).

3.2. Machine learning-based diagnostic value of various models in DFU prediction

Recognizing that combining multiple characteristic variables is more effective than relying on a single indicator in DFU diagnosis, this study aimed to assess the diagnostic capabilities of different combinations of characteristic indicators using machine learning techniques. We began by examining the two-dimensional distribution of various indicators in the training dataset and validation sets using t-SNE diagrams which represent the potential of various indicator combinations in the diagnosis process. Further, we analyzed the correlations among the indicators in the training set. The results confirmed that there were significant positive or negative correlations between some indicators, thus indicating a combined or antagonistic effect of these indicators in the diagnostic application. The details were shown in Figs. 2 and 3.

3.3. Performance of the predictive models

3.3.1. The first layer of machine learning

The average cross-validation performance of ML Layer 1 models were demonstrated in Table 4. Among the ML-1 prediction models with average MC cross-validation performance, the results from the confusion matrix analysis showed that all classifiers had a high test accuracy (74.0%–88.0 %), while only small discrepancy existed in the occurrence prediction rates of all classifiers (21%–29 %, Table 4). From Tables 4 and it could be seen that the RF model among all classifiers had the highest SNS (86 %), SPC (89 %), PPV (89 %), NPV (87 %), ACC (88 %) and OCC (29 %) values. The SVM model had lower SNS (77 %), SPC (71 %), PPV (73 %), NPV (76 %), and ACC (74 %) values than the other algorithm models, while its OCC value (26 %) is only lower than that of RF model. In the remaining models, the SNS, SPC, PPV, NPV, and ACC values were all between RF and SVM, but the BYS model had the lowest OCC value (21 %), and there existed large differences between SNS and PPV values, SPC and NPV values in this model. Overall, the percentage contribution of BYS occurrences is the smallest, thus it was not applied in ML-2.

3.3.2. The second layer of machine learning

The average cross-validation performance of ML Layer 1 models were demonstrated in Table 5. Three analystic methods (MGWC, RF and SVM) were screened by MC cross-validation. After two iterations of validation, all the three models of predicting the DFU occurrence showed a good ACC value (76%–87 %). Similar to the results in ML-1, RF model had the highest SNS (87 %), SPC (87 %), PPV (88 %)), NPV (88 %), and ACC (87 %) values, suggesting a better predictive performance than the other models. The evaluation parameters for SVM model were slightly higher than those in ML-1, but its SNS (79 %), SPC (73 %), PPV (75 %), NPV (79 %), and ACC (76 %) values were all lower than the other models. Besides, the evaluation parameters of MGWC were between those of RF and SVM models. In all prediction models, the OCC values were all equal to 33 %. Overall, The proportion weights of each algorithm model in cross-validation are equal. Despite the differences in predictive performance among the models, their performance is still acceptable.

3.3.3. The final predictive model

The accuracy and range of the parameters in the AutoML prediction model constructed based on the above three methods were presented in Table 6. Overall, the AutoML model had an excellent diagnostic performance in differentiating the risk of DFU occurrence. The results of ROC analysis showed that in the AutoML model the AUC of DFU occurrence was 88.48 % (ranging between 74.44 and 97.83 %), corresponding to SNS, SPC, ACC, PPV and NPV values of 87.23 (ranging between 63.33 and 100.00 %), 87.43 (ranging between 70.00 and 100.00 %), 87.33 (ranging between 76.66 and 95.00 %), 87.69 (ranging between 75.00 and 100.00 %), and 87.70

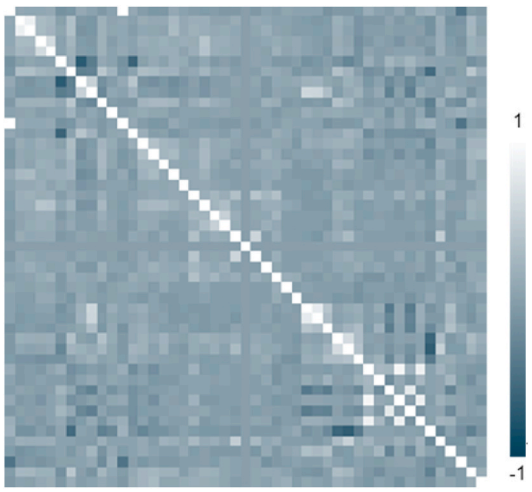


Fig. 2. Correlation value 1 and -1 mean a 100 % linear and inverse linear relationship between two features, respectively. Feature pairs with near 0 correlation value are considered non-redundant.

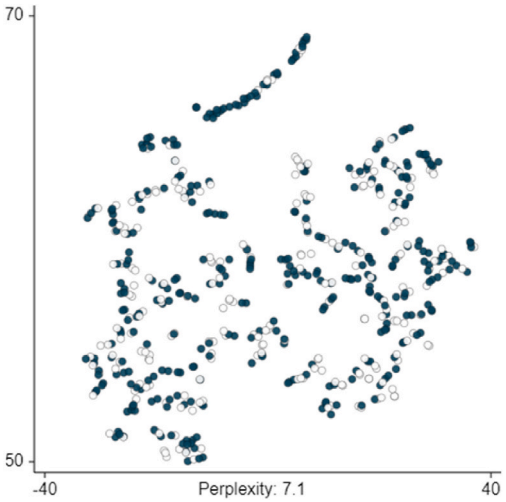


Fig. 3. T-distributed stochastic neighbor embedding (t-SNE) view of the data. Samples are colored by respective label outcomes.

Table 4
The average performance (%) of the predictive models in the first machine learning layer (ML-1) was determined using confusion matrix analysis across all Monte Carlo (MC) folds.

	SNS	SPC	PPV	NPV	ACC	OCC
BYS	84	71	75	83	77	21
MGWC	82	74	77	81	78	24
RF	86	89	89	87	88	29
SVM	77	71	73	76	74	26

Performance and occurrence values are in percentages.
BYS: Bayesian Classifier; MGWC: Multi-Gaussian Weighted Classifier; RF: Random Forest Classifier; SVM: Support Vector Machine Classifier; SNS: Sensitivity; SPC: Specificity; PPV: Positive Predictive Value; NPV: Negative Predictive Value; ACC: Accuracy; OCC: Occurrence.

Table 5
The average performance (%) of the predictive models in the second machine learning layer (ML-2) was determined using confusion matrix analysis across all Monte Carlo (MC) folds.

	SNS	SPC	PPV	NPV	ACC	OCC
MGWC	82	81	82	82	81	33
RF	87	87	88	88	87	33
SVM	79	73	75	79	76	33

Performance and occurrence values are in percentages.
MGWC: Multi-Gaussian Weighted Classifier; RF: Random Forest Classifier; SVM: Support Vector Machine Classifier; SNS: Sensitivity; SPC: Specificity; PPV: Positive Predictive Value; NPV: Negative Predictive Value; ACC: Accuracy; OCC: Occurrence.

Table 6
Performance Monte Carlo (MC) cross-validation performance of the established model scheme throughout the performance of the top-layer prediction model.

	Min	LQ	Median	UQ	Max	Mean	Dev
SNS	63.33	83.33	86.66	93.33	100	87.23	5.72
SPC	70.00	83.33	86.66	90.00	100	87.43	4.71
PPV	75.00	84.18	87.87	90.55	100	87.69	4.09
NPV	71.79	84.12	87.50	92.29	100	87.70	4.89
ACC	76.66	85.00	88.33	90.00	95.00	87.33	3.59
AUC	74.44	84.83	88.97	92.00	97.83	88.48	3.55

The performance values were determined using confusion matrix analysis across all Monte Carlo (MC) folds and were presented as percentages.
MGWC: Multi-Gaussian Weighted Classifier; RF: Random Forest Classifier; SVM: Support Vector Machine Classifier; SNS: Sensitivity; SPC: Specificity; PPV: Positive Predictive Value; NPV: Negative Predictive Value; ACC: Accuracy; AUC: Area Under the Receiver Operator Characteristics Curve. LQ: Lower quartile; UQ: Upper Quartile; Dev: Mean Deviation.

(ranging between 71.79 and 100.00 %), respectively (Table 6, Figs. 4 and 5).

In Fig. 6and Table 7, the importance ranking of the variables according to the AutoML prediction model was shown. In the contribution ranking, the patient's peripheral nerve symptoms such as pain, numbness, and sharp burning sensation were significantly the highest (16.3 %), followed by posterior tibial artery pulsation (6.12 %) and foot callus (5.57 %). The predictive contribution of the top 10 variables accounted for 52.2 %, with peripheral nerve symptoms of 23.7 %, peripheral artery assessment of 10.5 %, foot deformity of 6 %, systemic vascular disease of 9.5 % and DM course of 2.6 %, all in accordance with the DFU risk rating scale [37]. By the AutoML prediction model, other risk factors for DFU were further validated, such as metabolic disorders (like hyperuricemia, fatty liver, ACR), smoking, age and others. In addition, we identified new important variables that were not found by previous risk prediction methods by the AutoML prediction model, including varicose veins of lower limbs, medical history of cerebral infarction, blood urea nitrogen, GFR and type of diabetes, but the differences in predictive contribution among these variables were slight.

3.3.4. External validation results

A total of 158 samples were finally included in the external validation dataset (originally consisting of 248 samples, with 9 samples removed due to lack of labels and 81 samples removed due to missing 3 or more features; for samples with fewer than 3 missing features). The model demonstrated robust performance on the external dataset. Specifically, the Random Forest (RF) classifier exhibited an AUC of 0.762 (95 % CI: 0.730–0.806), sensitivity of 0.412 (95 % CI: 0.111–0.700), and specificity of 0.885 (95 % CI: 0.757–0.992). The Support Vector Machine (SVM) classifier achieved an AUC of 0.762 (95 % CI: 0.731–0.800), with a sensitivity of 0.477 (95 % CI: 0.117–0.700), and a specificity of 0.885 (95 % CI: 0.764–0.992). The Logistic Regression (LR) model showed an AUC of 0.739 (95 % CI: 0.729–0.746), sensitivity of 0.695 (95 % CI: 0.680–0.700), and specificity of 0.780 (95 % CI: 0.751–0.803). The details showed in Table 8 and Fig. 7.

4. Discussion

Although treatment strategies have been improving to better manage DFU in recent years, the healing and control of DFU still remains challenging because of the increasing DFU incidence and unsatisfying therapeutic effect. Early diagnosis and prevention keep extremely critical for the care of DFU [62]. However, at present, there are a variety of screening methods for DFU, while their promotion is limited by some accompanying shortcomings. For example, the time-consuming electrophysiological examination of lower extremity has a poor effect on subclinical patients [63], and QST may have high sensitivity but poor specificity [64]. On the other hand, more and more evidence shows that a single indicator is often unable to make an effective and accurate diagnosis, and the combined application of multiple indicators will become the trend [65], followed by more and more miscellaneous information input.

It was found that 68 % of the cases misdiagnosed as DFU turned out to be melanoma, 14 % Kaposi's sarcoma, 11 % squamous cell carcinoma, 4 % diffuse B-cell lymphoma, and 4 % mantle cell lymphoma. An older patient (age >65 years) was 145 % more likely to have such suspicions than a younger patient [66].

For medical staff, it will be more difficult to mine useful information from increasingly complex data [67]. Therefore, there is a need to develop convenient and intuitive tools to inspect existing variables, evaluate and include emerging variables, in order to support clinical transformation and practice to improve DFU screening.

AutoML aims to facilitate non-experts to apply machine learning models and techniques to solve specialized problems, create user interfaces and visualizations that maximize the value of every indicator, thus optimizing the efficient utilization of test data [68].In the early screening of DFU, predictive models have been increasingly applied [69,70], Numerous models have been proposed for differentiating DFU from healthy skin and predicting DFU progression. Emerging evidence suggests that predictive models can support clinicians in making diagnostic and therapeutic decisions, ultimately contributing to improved outcomes for patients with DFU [71], yet they remain in the exploratory phase and have not been standardized. Therefore, this study aims to contribute to the

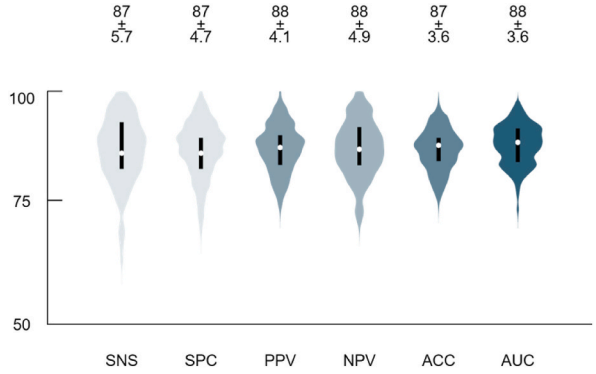


Fig. 4. Box-plot-based Monte Carlo (MC) cross-validation performance of the established model scheme throughout the performance of the top-layer prediction model. Performance values were determined by confusion matrix analysis across all MC folds and presented by percentages. SNS: Sensitivity; SPC: Specificity; PPV: Positive Predictive Value; NPV: Negative Predictive Value; ACC: Accuracy; AUC: Area Under the Receiver Operator Characteristics Curve.

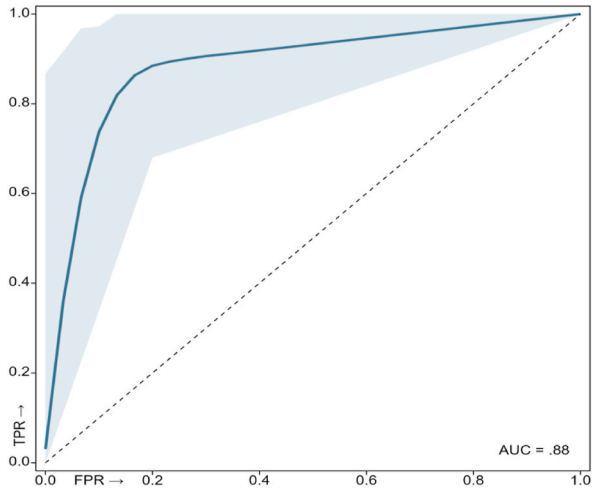


Fig. 5. The mean Receiver Operating Characteristic (ROC) curve determined by cross-validation for the built models. The mean ROC curve is represented by the thick blue line, while the light blue shaded area indicates the range of all 100 ROC curves generated across the validation folds. The dashed diagonal line is used as a reference for comparison, representing the AUC for random-guessing. AUC: Area Under the Receiver Operator Characteristics Curve (mean of all 100 AUCs); FPR: False Positive Rate; TPR: True Positive Rate.

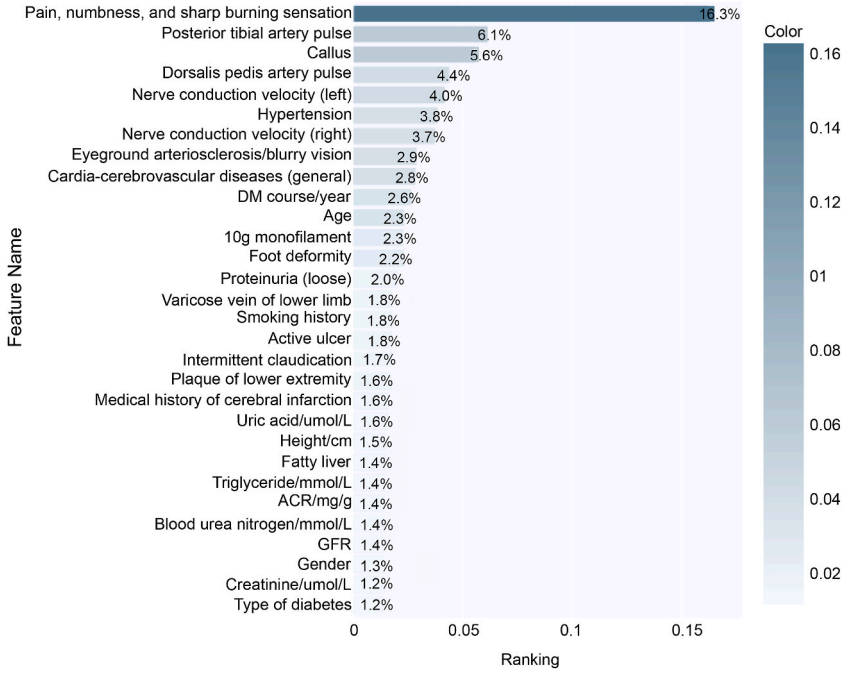


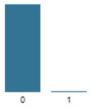
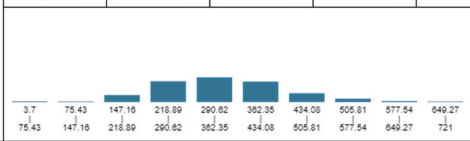
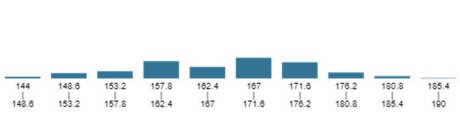
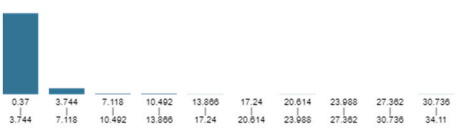
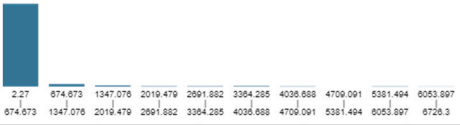
Fig. 6. The importance ranking of features.

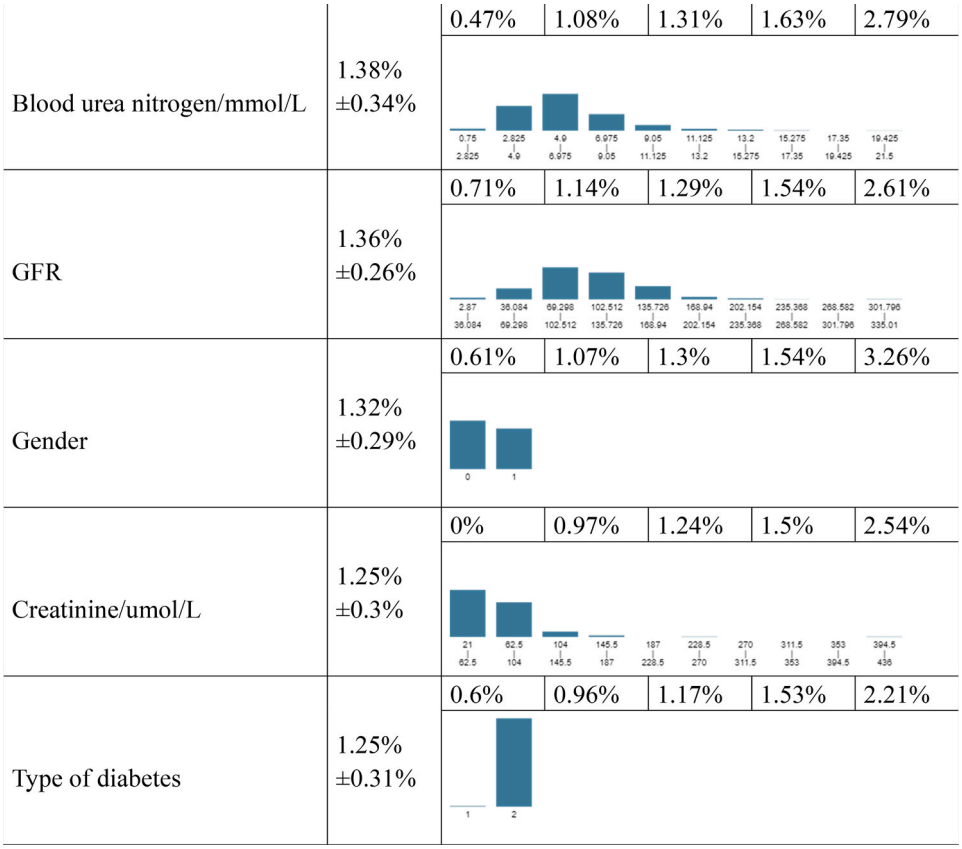
standardization of DFU screening by constructing relevant predictive models. In this study, a decision tool to screen DFU risk was successfully developed by AutoML. In ML-1, BYS model is excluded for ML-2 due to its poor performance, and this could be attributed to the fact that BYS is a probabilistic classifier that utilizes Bayes' theorem and assumes strong independence among the features during the modeling process. However, in the practical operation, features are more or less correlated, in which case the predictive performance of BYS may be greatly affected [49]. We have also noted that previous studies have primarily focused on employing traditional methods such as logistic regression or Cox proportional hazard regression analysis to examine the screening of risk factors. However, given the complex nature of DFU, its severity should be comprehensively determined by considering multiple factors, such as sociodemographic characteristics, anthropometric and laboratory test data. Since traditional statistical methods are limited by the linear relationship between variables, they may omit the complex relationship among many other important nonlinear interacting factors, thus weakening the explanatory power on results [72,73]. These findings suggest that the model's performance is influenced by the

Table 7
The features that were selected and their ranks, as determined by Smart Redundancy Reduction (SRR - seen in Table 1) across the Monte Carlo (MC) folds, are presented along with their corresponding value distributions.

Feature Name	Mean Ranking ±Dev	Min	LQ	Median	UQ	Max
		Feature Value Histogram				
Pain, numbness, and sharp burning sensation	16.3% ±1.86%	12.41%	14.54%	16.23%	17.68%	24.28%
Posterior tibial artery pulse	6.12% ±0.64%	4.45%	5.51%	6.12%	6.59%	8.79%
Callus	5.57% ±0.72%	3.74%	5.04%	5.44%	6.2%	8.9%
Dorsalis pedis artery pulse	4.39% ±0.53%	3.07%	3.86%	4.33%	4.85%	6.59%
Nerve conduction velocity (left)	3.99% ±0.58%	2.34%	3.55%	3.88%	4.42%	6.34%
Hypertension	3.83% ±0.52%	2.11%	3.38%	3.76%	4.26%	5.96%
		2.45%	3.18%	3.79%	4.13%	5.03%
Nerve conduction velocity (right)	3.70% ±0.51%					
Eyeground arteriosclerosis/blurry vision	2.89% ±0.51%	1.39%	2.44%	2.85%	3.2%	4.66%
Cardia-cerebrovascular diseases (general)	2.83% ±0.44%	1.68%	2.47%	2.72%	3.13%	5.01%

DM course/year	2.57% ±0.41%	1.28%	2.19%	2.49%	2.96%	4.9%
Age	2.30% ±0.33%	1.34%	2%	2.26%	2.53%	3.5%
Semmes-Weinstein monofilament test	2.27% ±0.29%	1.44%	2.01%	2.21%	2.55%	3.59%
Foot deformity	2.25% ±0.4%	1.11%	1.88%	2.18%	2.47%	5.33%
Proteinuria (loose)	1.99% ±0.38%	1.1%	1.67%	1.92%	2.32%	3.38%
Varicose vein of lower limb	1.76% ±0.41%	0.72%	1.36%	1.66%	2.08%	3.26%
Smoking history	1.75% ±0.43%	0.77%	1.34%	1.65%	2.02%	3.7%
Active ulcer	1.75% ±0.44%	0%	1.37%	1.67%	2.06%	3.46%

Intermittent claudication	1.68% ±0.35%	0%	1.39%	1.65%	2.01%	3.69%
						
Plaque of lower extremity	1.64% ±0.32%	0.76%	1.35%	1.61%	1.91%	2.98%
						
Medical history of cerebral infarction	1.59% ±0.33%	0.8%	1.29%	1.52%	1.84%	2.87%
						
Uric acid/umol/L	1.57% ±0.36%	0.63%	1.25%	1.52%	1.82%	2.7%
						
Height/cm	1.47% ±0.42%	0%	1.07%	1.36%	1.72%	3.81%
						
Fatty liver	1.45% ±0.41%	0.49%	1.1%	1.31%	1.61%	6.54%
						
		0.72%	1.13%	1.41%	1.71%	2.9%
Triglyceride/mmol/L	1.43% ±0.34%					
ACR/mg/g	1.41% ±0.33%	0.38%	1.15%	1.36%	1.7%	2.34%
						



The ranks indicate the relative significance of the selected features in building the model, with features being ordered accordingly. The rank values are expressed as percentages.
LQ: Lower quartile; UQ: Upper Quartile; Dev: Mean Deviation.

Table 8
The average performance (%) of the External Validation of the Predictive Model.

	SNS	SPC	PPV	NPV	ACC	AUC
SVM	0.477 (0.117–0.700)	0.885 (0.764–0.992)	0.658 (0.501–0.846)	0.841 (0.767–0.885)	0.782 (0.747–0.820)	0.762 (0.731–0.800)
RF	0.412 (0.111–0.700)	0.885 (0.757–0.992)	0.664 (0.495–0.856)	0.827 (0.766–0.886)	0.765 (0.743–0.779)	0.762 (0.730–0.806)
LR	0.695 (0.680–0.700)	0.780 (0.751–0.803)	0.518 (0.488–0.540)	0.883 (0.880–0.886)	0.758 (0.738–0.772)	0.739 (0.729–0.746)

SVM: Support Vector Machine Classifier; RF: Random Forest Classifier; LR: logistic regression; SNS: Sensitivity; SPC: Specificity; PPV: Positive Predictive Value; NPV: Negative Predictive Value; ACC: Accuracy; AUC: Area Under the Curve.

characteristics of the dataset, underscoring the importance of selecting an appropriate model based on the data type. After iterative verification in this study, the predictive performance of each model was found to be comparable. The final model, which was constructed using a combination of SVM, RF, and MGWC, exhibited excellent predictive performance with an AUC of 88.48 (ranging from 74.44 to 97.83). Notably, this performance was superior to that of traditional risk score models [38]. In order to verify the applicability and reliability of the prediction model, external datasets were used in this study. The results of external datasets demonstrated that the AutoML model maintained a high level of predictive accuracy in the external dataset. The AUC score of 0.762 indicated a strong discriminatory capability of the model. Furthermore, the Sensitivity and Specificity values provided insights into the model's ability to correctly identify both DFU cases and non-cases, respectively. This finding aligns with the research conducted by Hong et al. [74], who highlighted the utility of machine learning in identifying individuals at high risk for DFU. Their work provides further support for our model, particularly in the context of DFU risk assessment tailored to specific patient populations.

This study reveals the top 10 key parameters prominent in the process of DFU screening, which is consistent with the emphasis of Chinese DFU risk screening [38]. The findings of this study indicate that when as many risk factors as possible are included, peripheral nerve symptoms, peripheral artery assessment and foot deformity remain the key screening items of DFU. As a program of DFU peripheral nerve screening, Semmes-Weinstein monofilament test is a widely-applied routine means [75]. Besides, intermittent

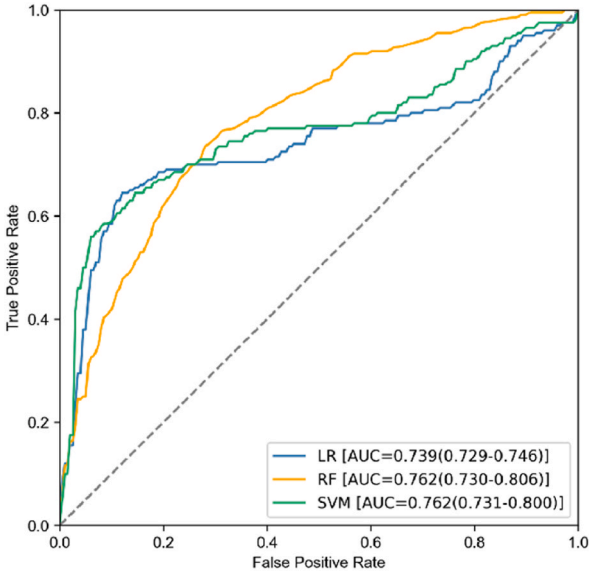


Fig. 7. The AUC of the predictive model.

claudication is a common symptom of peripheral artery injury [76]. However, they are not in a high ranking of the prediction contribution, possibly because Although there is a high correlation among some project features, the diagnostic sensitivity of the diagnostic items themselves may not be high, or they may be less common compared to high-ranking projects with high predictive contributions [65]. That also means when developing a screening tool for the DFU, the two items can replace to some extent others in the same category. It is worth noting that in the traditional DFU risk screening score [11,38], there is no assessment of systemic vascular disease, while in this study, this item accounts for 9.5 % among the top 10 items (Hypertension, Eyeground arteriosclerosis/blurry vision, Cardia-cerebrovascular diseases (general)) in the prediction contribution ranking, higher than that of the foot deformity (6 %). It indicates that in addition to local foot features, the overall vascular function assessment of patients should also be a fully considered part of the clinical screening score.

Consistent with previous studies [77], the results of this study further demonstrate that metabolic disorders (such as hyperuricemia, fatty liver and ACR), abnormal renal function, smoking, age, height and so on are all important risk factors of DFU. Interestingly, we identify some new risk factors (GFR, varicose veins of lower extremity, medical history of cerebral infarction, blood urea nitrogen and type of diabetes) that had not been detected in previous risk screening by AutoML, suggesting that machine learning techniques have unique advantages in identifying new and important risk factors in epidemiological studies. It is also noted that the remaining 20 projects in this study account for 47.8 % of the total prediction contribution, and the differences between these projects are small, which means that each project has a similar degree of correlation to the DFU.

Some of the clinical findings mentioned, such as pain, numbness, sharp burning sensation, history of cerebral infarction, blood urea nitrogen, and GFR, have been recognized in medicine for many decades as associated with diabetes. These factors are indeed considered well-known long-term consequences of diabetes and are commonly associated with the risk of DFU or/their proxies.

While the individual clinical findings may not be new, it is still important to highlight their association with the risk of DFU. The research by Tuglo et al. [78]. presents a concerning phenomenon: individuals with diabetes mellitus (DM), despite possessing good knowledge about DFU care, exhibit suboptimal DFU management practices. This discrepancy arises due to various factors, including busy lifestyles, lack of encouragement and motivation, or insufficient prioritization of health risks. By emphasizing these factors, healthcare professionals can help raise awareness about the importance of early detection, monitoring, and management of these conditions in individuals with diabetes. Recognizing the impact of these factors on DFU risk can guide healthcare providers in tailoring interventions and treatments to reduce the occurrence and severity of DFU.

It is crucial to consider the multifactorial nature of DFU and the unique combination of risk factors that each patient may have.

Taking the existing problems at the present stage into consideration, we make sure that the direction of breakthrough in the future should be how to complement the new indicators with traditional indicators for screening and how to maximize the integration of the traditional indicators to improve the DFU risk screening. What's more, economic and human cost is an inevitable problem. In this study, the finally selected 30 items involve the general characteristics of patients, physical data, laboratory and imaging data. How to screen and manage DFU at the system level to strictly control the cost should be carefully considered in combination with the outstanding contribution items and the individual conditions of patients.

Although the accuracy of this prediction model is improved compared with the Chinese version of DFU risk screening tool [38] by AutoML training model, this study has certain limitations. Firstly, the performance of the predictive model based on AutoML is influenced by the sample size, which was relatively small in this study. The sample size of 566 cases used for the modeling is indeed a relatively small sample, which may limit the generalizability of the results. Therefore, it is imperative to further validate and refine the

model's performance through large-scale training to ensure its clinical applicability. Secondly, this study lacks patients with end-stage renal failure, diabetic foot amputation or ulcer history and so on. In the future studies, such kinds of patients should be included to verify the value of arterial pulsation test in DFU risk prediction by combining objective indicators like ankle-brachial index (ABI). Thirdly, the internal complexity of AutoML makes its internal logic difficult to understand. Additionally, the lack of validation with fresh data means that the model's performance and accuracy could vary when applied to new or different datasets. However, given the absence of missing values in the data batches of this study, and the external validation offered compelling evidence of the model's robustness and generalizability, Despite potential variations in clinical presentation and patient demographics, the model performed consistently, suggesting its potential for widespread clinical application. we are optimistic about the performance of the final machine learning model.

Therefore, it is sometimes difficult to understand the influence of complex interactions among multiple internal variables on the results, which may undermine the confidence of external applications. It's important to approach the findings with caution and recognize that further research and validation are needed to confirm the results and assess their applicability in different populations or settings. Replication of the study with larger sample sizes and validation using independent datasets would enhance the reliability and robustness of the findings.

In the meantime, it would be prudent to consider these results as exploratory and hypothesis-generating rather than conclusive evidence. This also reminds us that further validation is needed in larger and more diverse populations to fully determine the reliability of this model in various healthcare settings. In conclusion, we confirm that on the basis of comprehensive information and abundant data, the predictive model established by AutoML is a practical and powerful tool for the screening and diagnosis of DFU occurrence, which can deeply understand the risk factors of diseases without assuming causality in advance and contribute to the early and accurate diagnosis, treatment and management of DFU patients to improve the quality of life.

5. Conclusions

The consequences of diabetic foot problems are significant, not only in terms of mortality, but also in terms of the cost of controlling and monitoring the disease. In order to prevent such complications, early detection and severity classification are crucial. It may be possible to develop easy-to-use solutions for early detection of disease by deploying machine learning in biomedical applications, not only do such solutions save medical professionals time, but they can also be beneficial for patients at home. They can use it in their homes, especially during pandemic times, where visits to healthcare services are preferred to be limited, avoiding stress on the healthcare system. Based on local foot examinations, the author of the paper proposed a new practical and powerful tool for screening, in which Peripheral nerve pain, posterior tibial artery, foot calluses and dorsal foot artery pulsation account for 32.38 % of the total weight. By combining the model of local examination, people with high-risk diabetes feet can be screened out as soon as possible.This content can also be used at home, allowing remote applications to be used to screen individuals at risk of DFU.

Funding

This study was supported by the Shanghai Municipal Education Commission-Gaoyuan Nursing Grant Support (Hlgy1810kyx), Xin Hua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine Nursing Development Program (HLXKRC2021004), Innovative Research Team of High-Level Local Universities in Shanghai (SHSMU-ZLCX20212401) and National Natural Science Foundation of China (72104102).

CRediT authorship contribution statement

Wang Xiaoling: Writing – original draft, Funding acquisition, Conceptualization. **Zhu Shengmei:** Writing – original draft, Conceptualization. **Wang BingQian:** Writing – original draft, Methodology. **Li Wen:** Data curation. **Gu Shuyan:** Software, Funding acquisition. **Chen Hanbei:** Data curation. **Qin Chenjie:** Writing – review & editing, Validation, Formal analysis. **Dai Yao:** Writing – review & editing, Project administration. **Li Jutang:** Writing – review & editing, Visualization, Validation, Project administration, Formal analysis.

Declaration of competing interest

We declare that we have no financial and personal relationships with other people ororganizations that can inappropriately influence our work, there is no professional orother personal interest of any nature or kind in any product, service and/or companythat could be construed as influencing the position presented in, or the review of,the manuscript entitled.

References

[1] L. Chen, et al., Global mortality of diabetic foot ulcer: a systematic review and meta-analysis of observational studies, Diabetes Obes. Metabol. 25 (1) (2023) 36–45, https://doi.org/10.1111/dom.14840.

[2] K. McDermott, et al., Etiology, epidemiology, and disparities in the burden of diabetic foot ulcers, Diabetes Care 46 (1) (2022) 209–221, https://doi.org/10.2337/dci22-0043.

[3] World-Health-Organization, Global report on diabetes. https://www.who.int/publications/i/item/9789241565257, 2016.

[4] Q. Guo, et al., Influencing factors for the recurrence of diabetic foot ulcers: a meta-analysis, Int. Wound J. 20 (5) (2023) 1762–1775, https://doi.org/10.1111/iwj.14017.
[5] Y. Jiang, et al., A cohort study of diabetic patients and diabetic foot ulceration patients in China, Wound Repair Regen. 23 (2) (2015) 222–230.
[6] A. Wang, et al., Guidelines on Multidisciplinary Approaches for the Prevention and Management of Diabetic Foot Disease, 2020 edition, Burns & Trauma, 2020, p. tkaa017, https://doi.org/10.1093/burnst/tkaa017.
[7] F. Bekele, L. Chelkeba, Amputation rate of diabetic foot ulcer and associated factors in diabetes mellitus patients admitted to Nekemte referral hospital, western Ethiopia: prospective observational study, J. Foot Ankle Res. 13 (1) (2020) 65.
[8] J.W. Walsh, et al., Association of diabetic foot ulcer and death in a population-based cohort from the United Kingdom, Diabet. Med. 33 (11) (2016) 1493–1498, https://doi.org/10.1111/dme.13054.
[9] J. Tulloch, R. Zamani, M. Akrami, Machine learning in the prevention, diagnosis and management of diabetic foot ulcers A systematic review, IEEE Access 8 (2020) 198977–199000, https://doi.org/10.1109/access.2020.3035327.
[10] A. Taghizadeh, et al., Mussel-inspired biomaterials: from chemistry to clinic, Bioengineering & Translational Medicine 7 (3) (2022) e10385, https://doi.org/10.1002/btm2.10385.
[11] Scottish-intercollegiate-Guidelines-network, Management of Diabetes A National Clinical Guideline, 2.20, 2022. https://www.sign.ac.uk/assets/sign116.pdf.
[12] K.K. Kaur, G. Allahbadia, M. Singh, The role of Diabetes Mellitus (both T1D and T2D) in the atherosclerosis development—a systematic review with part of inflammation along with altered glucose and lipid metabolism for forming therapeutic aproaches, EC Diabetes Metab. Res 4 (2020) 27–43.
[13] G. Daryabor, et al., The effects of type 2 diabetes mellitus on organ metabolism and the immune system, Front. Immunol. 11 (2020) 1582, https://doi.org/10.3389/fimmu.2020.01582.
[14] J. Ilonen, J. Lempainen, R. Veijola, The heterogeneous pathogenesis of type 1 diabetes mellitus, Nat. Rev. Endocrinol. 15 (11) (2019) 635–650, https://doi.org/10.1038/s41574-019-0254-y.
[15] A.J. Boulton, et al., Diagnosis and management of diabetic foot infections, Compendia 2020 (1) (2020), https://doi.org/10.2337/db2020-01.
[16] C.E. Hazenberg, et al., Telehealth and telemedicine applications for the diabetic foot: a systematic review, Diabetes/metabolism research and reviews 36 (3) (2020) e3247, https://doi.org/10.1002/dmrr.3247.
[17] M.H. Yap, et al., Deep learning in diabetic foot ulcers detection: a comprehensive evaluation, Comput. Biol. Med. 135 (2021) 104596.
[18] P. Xie, et al., Mortality prediction in patients with hyperglycaemic crisis using explainable machine learning: a prospective, multicentre study based on tertiary hospitals, Diabetol. Metab. Syndrome 15 (1) (2023) 44, https://doi.org/10.1186/s13098-023-01020-1.
[19] R. Naemi, et al., Predicting the risk of future diabetic foot ulcer occurrence: a prospective cohort study of patients with diabetes in Tanzania, BMJ Open Diabetes Research & Care 8 (1) (2020) e001122.
[20] X. Yuan, et al., A multi-centered study in diabetic at-risk foot screening and risk factor analysis in Beijing, Chinese Journal of diabetes 28 (7) (2020) 486–491.
[21] Q. Lu, et al., Risk factors for major amputation in diabetic foot ulcer patients, Diabetes, Metabolic Syndrome and Obesity 14 (null) (2021) 2019–2027.
[22] L. Wang, et al., Prediction of type 2 diabetes risk and its effect evaluation based on the XGBoost model, Healthcare 8 (3) (2020) 247.
[23] J.D. Brocklehurst, The validity and reliability of the SINBAD classification system for diabetic foot ulcers, Adv. Skin Wound Care 36 (11) (2023) 1–5.
[24] P. Normahani, et al., Cost-effectiveness of diagnostic tools to establish the presence of peripheral arterial disease in people with diabetes, Ann. Surg. 277 (1) (2023) e184–e191.
[25] T. Howard, R. Ahluwalia, N. Papanas, The advent of artificial intelligence in diabetic foot medicine: a new horizon, a new order, or a False dawn? Int. J. Low. Extrem. Wounds 22 (4) (2023) 635–640, https://doi.org/10.1177/15347346211041866.
[26] Z. Xueyan, et al., Regional differences of quality medical resources in China based on different scales, Econ. Geogr. 40 (7) (2020) 22–31, https://doi.org/10.15957/j.cnki.jjdl.2020.07.003.
[27] A. Rajkomar, J. Dean, I. Kohane, Machine learning in medicine, N. Engl. J. Med. 380 (14) (2019) 1347–1358, https://doi.org/10.1056/NEJMra1814259.
[28] A. Nomura, et al., Artificial intelligence in current diabetes management and prediction, Curr. Diabetes Rep. 21 (12) (2021) 1–6, https://doi.org/10.1007/s11892-021-01423-2.
[29] M.A. Makroum, et al., Machine learning and smart devices for diabetes management: systematic review, Sensors 22 (5) (2022) 1843.
[30] M.B. Alazzam, et al., Machine learning implementation of a diabetic patient monitoring system using interactive E-app, Comput. Intell. Neurosci. 2021 (2021), https://doi.org/10.1155/2021/5759184.
[31] N. Sneha, T. Gangil, Analysis of diabetes mellitus for early prediction using optimal features selection, Journal of Big data 6 (1) (2019) 1–19, https://doi.org/10.1186/s40537-019-0175-6.
[32] F. Arcadu, et al., Deep learning algorithm predicts diabetic retinopathy progression in individual patients, NPJ digital medicine 2 (1) (2019) 1–9, https://doi.org/10.1038/s41746-019-0172-3.
[33] A. Khandakar, et al., A machine learning model for early detection of diabetic foot using thermogram images, Comput. Biol. Med. 137 (2021) 104838, https://doi.org/10.1016/j.compbiomed.2021.104838.
[34] A. Khandakar, et al., A novel machine learning approach for severity classification of diabetic foot complications using thermogram images, Sensors 22 (11) (2022) 4249.
[35] S.K. Das, P. Roy, A.K. Mishra, DFU_SPNet: a stacked parallel convolution layers based CNN to improve Diabetic Foot Ulcer classification, ICT Express 8 (2) (2022) 271–275, https://doi.org/10.1016/j.icte.2021.08.022.
[36] S. Stefanopoulos, et al., Machine learning prediction of diabetic foot ulcers in the inpatient population, Vascular (2021) 17085381211040984, https://doi.org/10.1177/17085381211040984.
[37] International-Working-Group-on-the-Diabetic-Foot, IWGDF Practical guidelines on the prevention and management of diabetic foot disease, 2023. https://iwgdfguidelines.org/wp-content/uploads/2023/07/IWGDF-2023-01-Practical-Guidelines.pdf.
[38] X.l. Wang, et al., Predictive test for Chinese version of the diabetes foot screening and risk stratification tool, Journal of Nursing Science 36 (17) (2021) 13–16, https://doi.org/10.3870/j.issn.1001-4152.2021.17.013.
[39] L. Papp, et al., Personalizing medicine through hybrid imaging and medical big data analysis, Frontiers in Physics 6 (2018) 51, https://doi.org/10.3389/fphy.2018.00051.
[40] A.A. Naeini, M. Babadi, S. Homayouni, Assessment of normalization techniques on the accuracy of hyperspectral data clustering. International archives of the photogrammetry, Remote Sensing & Spatial Information Sciences 42 (2017), https://doi.org/10.5194/isprs-archives-xlii-4-w4-27-2017.
[41] W.I.D. Mining, Data mining: concepts and techniques, Morgan Kaufinann 10 (2006) 559–569, https://doi.org/10.1115/1.2123107.
[42] D.K. Choubey, et al., Classification of diabetes by kernel based SVM with PSO, Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science) 14 (4) (2021) 1242–1255, https://doi.org/10.2174/2213275912666190716094836.
[43] P. Pujari, Classification of pima Indian diabetes dataset using support vector machine with polynomial kernel, in: Deep Learning, Machine Learning and IoT in Biomedical and Health Informatics, CRC Press, 2022, pp. 55–67.
[44] S. Cr, Kernel functions for machine learning applications. http://crsouza.blogspot.com/2010/03/kernel-functions-for-machine-learning.html, 2020.
[45] C.C. Olisah, L. Smith, M. Smith, Diabetes mellitus prediction and diagnosis from a data preprocessing and machine learning perspective, Comput. Methods Progr. Biomed. 220 (2022) 106773, https://doi.org/10.1016/j.cmpb.2022.106773.
[46] Y.-T. Wu, et al., Early prediction of gestational diabetes mellitus in the Chinese population via advanced machine learning, The Journal of Clinical Endocrinology & Metabolism 106 (3) (2021) e1191–e1205, https://doi.org/10.1210/clinem/dgaa899.
[47] L. Papp, et al., Supervised machine learning enables non-invasive lesion characterization in primary prostate cancer with [68Ga]Ga-PSMA-11 PET/MRI, Eur. J. Nucl. Med. Mol. Imaging 48 (2020) 1795–1805.
[48] H. Zhang, The optimality of naive Bayes, AAAI 1 (2) (2004) 3.
[49] I. Rish, An empirical study of the naive Bayes classifier, in: IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence, 2001.

[50] B. Zong, et al., Deep autoencoding Gaussian mixture model for unsupervised anomaly detection, in: International Conference on Learning Representations, 2018.
[51] C. Viroli, G.J. McLachlan, Deep Gaussian mixture models, Stat. Comput. 29 (2019) 43–51, https://doi.org/10.1007/s11222-017-9793-z.
[52] D.A. Reynolds, Gaussian mixture models, Encyclopedia of biometrics 741 (2009) 659–663, https://doi.org/10.1007/springerreference_70943.
[53] C. Cortes, V. Vapnik, Support-vector networks, Mach. Learn. 20 (3) (1995) 273–297, https://doi.org/10.1007/BF00994018.
[54] W.S. Noble, What is a support vector machine? Nat. Biotechnol. 24 (12) (2006) 1565–1567, https://doi.org/10.1038/nbt1206-1565.
[55] V. Svetnik, et al., Random forest: a classification and regression tool for compound classification and QSAR modeling, J. Chem. Inf. Comput. Sci. 43 (6) (2003) 1947–1958, https://doi.org/10.1021/ci034160g.
[56] V. Podgorelec, et al., Decision trees: an overview and their use in medicine, J. Med. Syst. 26 (5) (2002) 445–463, https://doi.org/10.1023/A:1016409317640.
[57] R.J. Marshall, The use of classification and regression trees in clinical epidemiology, Journal of clinical epidemiology 54 (6) (2001) 603–609, https://doi.org/10.1016/S0895-4356(00)00344-9.
[58] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32.
[59] M.J. Van der Laan, E.C. Polley, A.E. Hubbard, Super learner, Stat. Appl. Genet. Mol. Biol. 6 (1) (2007), https://doi.org/10.2202/1544-6115.1309.
[60] M. Zhao, et al., Multi-lesion radiomics of PET/CT for non-invasive survival stratification and histologic tumor risk profiling in patients with lung adenocarcinoma, Eur. Radiol. 32 (10) (2022) 7056–7067.
[61] S.V. Stehman, Selecting and interpreting measures of thematic classification accuracy, Rem. Sens. Environ. 62 (1) (1997) 77–89, https://doi.org/10.1016/S0034-4257(97)00083-7.
[62] M.H. Alshayeji, S. ChandraBhasi Sindhu, S.e. Abed, Early detection of diabetic foot ulcers from thermal images using the bag of features technique, Biomed. Signal Process Control 79 (2023) 104143, https://doi.org/10.1016/j.bspc.2022.104143.
[63] A.S. Dikici, et al., Evaluation of the tibial nerve with shear-wave elastography: a potential sonographic method for the diagnosis of diabetic peripheral neuropathy, Radiology 282 (2) (2017) 494–501, https://doi.org/10.1148/radiol.2016160135.
[64] D. Wang, et al., MR T2 value of the tibial nerve can be used as a potential non-invasive and quantitative biomarker for the diagnosis of diabetic peripheral neuropathy, Eur. Radiol. 28 (3) (2018) 1234–1241, https://doi.org/10.1007/s00330-017-5043-1.
[65] Y. Luo, et al., Machine learning based on routine laboratory indicators promoting the discrimination between active tuberculosis and latent tuberculosis infection, J. Infect. 84 (5) (2022) 648–657, https://doi.org/10.1016/j.jinf.2021.12.046.
[66] A.V. Lyundup, et al., Misdiagnosis of diabetic foot ulcer in patients with undiagnosed skin malignancies, Int. Wound J. 19 (4) (2022) 871–887, https://doi.org/10.1111/iwj.13688.
[67] J. Waring, C. Lindvall, R. Umeton, Automated machine learning: review of the state-of-the-art and opportunities for healthcare, Artif. Intell. Med. 104 (2020) 101822, https://doi.org/10.1016/j.artmed.2020.101822.
[68] I. Guyon, et al., Analysis of the AutoML challenge series, Automated Machine Learning 177 (2019).
[69] Y. Chen, J. Zhuang, C. Yang, Development of a major amputation prediction model and nomogram in patients with diabetic foot, Postgrad. Med. (2024), https://doi.org/10.1093/postmj/qgae087.
[70] J. Lv, et al., Development and validation of a risk prediction model for foot ulcers in diabetic patients, J. Diabetes Res. 2023 (1) (2023) 1199885, https://doi.org/10.1155/2023/1199885.
[71] T. Weatherall, et al., The impact of machine learning on the prediction of diabetic foot ulcers – a systematic review, J. Tissue Viability (2024), https://doi.org/10.1016/j.jtv.2024.07.004.
[72] X. Wu, et al., Data mining with big data, IEEE Trans. Knowl. Data Eng. 26 (1) (2013) 97–107, https://doi.org/10.1109/TKDE.2013.109.
[73] S. Hengl, et al., Data-based identifiability analysis of non-linear dynamical models, Bioinformatics 23 (19) (2007) 2612–2618, https://doi.org/10.1093/bioinformatics/btm382.
[74] S. Hong, et al., Personalized prediction of diabetic foot ulcer recurrence in elderly individuals using machine learning paradigms, Technol. Health Care 32 (2024) 265–276, https://doi.org/10.3233/THC-248023.
[75] F. Wang, et al., Diagnostic accuracy of monofilament tests for detecting diabetic peripheral neuropathy: a systematic review and meta-analysis, J. Diabetes Res. (2017) 2017, https://doi.org/10.1155/2017/8787261.
[76] A. Andrade-Lima, et al., Walking training improves systemic and local pathophysiological processes in intermittent claudication, Eur. J. Vasc. Endovasc. Surg. 61 (6) (2021) 954–963, https://doi.org/10.1016/j.ejvs.2021.02.022.
[77] Chinese Diabetes Society, Chinese Society of Infectious Diseases, Chinese Society for Tissue Repair, Regeneration, Chinese guideline on prevention and management of diabetic foot (2019 edition)(I), Chin. J. Diabetes Mellitus 11 (2) (2019) 92–108, https://doi.org/10.3760/cma.j.issn.1674-5809.2019.02.004.
[78] L.S. Tuglo, et al., Knowledge and practice of diabetic foot care and the prevalence of diabetic foot ulcers among diabetic patients of selected hospitals in the Volta Region, Ghana, Int. Wound J. 19 (3) (2022) 601–614, https://doi.org/10.1111/iwj.13656.