scientific reports

Check for updates

OPEN

A machine learning-based framework for predicting metabolic syndrome using serum liver function tests and high-sensitivity C-reactive protein

Bahareh Behkamal[1,2], Fatemeh Asgharian Rezae[2], Amin Mansoori[3]✉, Rana Kolahi Ahari[4,5], Sobhan Mahmoudi Shamsabad[6], Mohammad Reza Esmaeilian[6], Gordon Ferns[7], Mohammad Reza Saberi[1,8]✉, Habibollah Esmaily[9,10]✉ & Majid Ghayour-Mobarhan[5,11]

Metabolic Syndrome (MetS) comprises a clustering of conditions that significantly increase the risk of heart disease, stroke, and diabetes. Timely detection and intervention are crucial in preventing severe health outcomes. In this study, we implemented a machine learning (ML)-based predictive framework to identify MetS using serum liver function tests—Alanine Transaminase (ALT), Aspartate Aminotransferase (AST), Direct Bilirubin (BIL.D), Total Bilirubin (BIL.T)—and high-sensitivity C-reactive protein (hs-CRP). The framework integrated diverse ML algorithms, including Linear Regression (LR), Decision Trees (DT), Support Vector Machine (SVM), Random Forest (RF), Balanced Bagging (BG), Gradient Boosting (GB), and Convolutional Neural Networks (CNNs). This framework is designed to develop a robust, scalable, and efficient predictive tool. We evaluated our approach on a large-scale cohort comprising 9,704 participants from the Mashhad Stroke and Heart Atherosclerotic Disorder (MASHAD) study, spanning 2010–2020. After preprocessing, a final dataset of 8,972 individuals (3,442 with MetS and 5,530 without) was used for model development and validation. Among the tested models, GB and CNN demonstrated superior performance, achieving specificity rates of 77% and 83%, respectively. The Gradient Boosting model achieved the lowest error rate of 27%, indicating robust predictive capability. Additionally, SHAP analysis identified hs-CRP, BIL.D, ALT, and sex as the most influential predictors of MetS. These findings suggest that leveraging liver function biomarkers and hs-CRP within an automated ML pipeline can facilitate early, non-invasive detection of MetS, supporting clinical decision-making and risk stratification efforts in healthcare systems.

Keywords Metabolic syndrome, Liver function tests, High-sensitivity C-reactive protein (hs-CRP), Machine learning

[1]Medicinal Chemistry Department, School of Pharmacy, Mashhad University of Medical Sciences, Mashhad 9177899191, Iran. [2]Student Research Committee, Mashhad University of medical sciences, Mashhad, Iran. [3]Department of Applied Mathematics, School of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad 9177948953, Iran. [4]Applied Biomedical Research Center, Mashhad University of Medical Sciences, Mashhad, Iran. [5]International UNESCO center for Health-Related Basic Sciences and Human Nutrition, Mashhad University of Medical Sciences, Mashhad, Iran. [6]Department of Computer Engineering, Faculty of Engineering, Ferdowsi University of Mashhad, Mashhad 9177948974, Iran. [7]Division of Medical Education, Brighton and Sussex Medical School, Brighton, UK. [8]Bioinformatics Research Group, Mashhad University of Medical Sciences, Mashhad 9177899191, Iran. [9]Department of Biostatistics, School of Health, Mashhad University of Medical Sciences, Mashhad, Iran. [10]Social Determinants of Health Research Center, Mashhad University of Medical Sciences, Mashhad, Iran. [11]Metabolic Syndrome Research Center, Mashhad University of Medical Sciences, Mashhad, Iran. ✉email: aminmansoori@um.ac.ir; saberiMR@mums.ac.ir; esmailyh@mums.ac.ir

Abbreviations

ML	Machine Learning
ANN	Artificial Neural network
LR	Linear Regression
DT	Decision Tree
SVM	Support Vector Machine
RF	Random Forests
GB	Gradient Boosting
BG	Balanced Bagging
CNNs	Convolutional Neural Networks
MASHAD	Mashhad Stroke and Heart Atherosclerotic Disorder
Hs-CRP	High-Sensitivity C-reactive Protein
MetS	Metabolic Syndrome
ALT	Alanine Aminotransferase
AST	Aspartate Aminotransferase
BIL.D	Direct Bilirubin
BIL.T	Total Bilirubin
CVD	Cardiovascular Disease
HTN	Hypertension
NAFLD	Non-Alcoholic Fatty Liver Disease
T2DM	Type 2 Diabetes Mellitus
BMI	Body Mass Index
AUC	Area Under Curve
ROC	Receiver Operator Characteristic
SHAP	SHapely Additive exPlanations
ROS	Random Over-Sampling
SMOTE	Synthetic Minority Oversampling Technique
RUS	Random Under-Sampling
TP	True Positives
TN	True Negatives
FN	False Negatives
FP	False Positives

Metabolic Syndrome (MetS) represents a significant public health concern due to its rising prevalence and substantial role in increasing morbidity and mortality associated with cardiovascular diseases and diabetes. MetS is a clustering of cardiovascular risk factors such as hypertension (HTN), hyperglycemia, abdominal obesity, and abnormal levels of blood lipids (increased triglycerides and low high-density lipoprotein (HDL) concentrations)[1–3]. Early detection and proactive management of metabolic syndrome can greatly reduce health risks and improve an individual's quality of life, making awareness and regular health screening important components of health care. MetS is typically diagnosed through a combination of clinical and biochemical parameters, which include waist circumference, lipid levels, blood pressure, fasting glucose levels. In this study, we leveraged underexplored biomarkers from liver function tests and inflammation indicator, high-sensitivity C-reactive protein (hs-CRP), to predict MetS within a cohort dataset. Liver enzymes and hs-CRP have been identified as potential markers due to their correlation with inflammation and liver fat accumulation, which are common in MetS patients. Leveraging these biomarkers for early detection of MetS can facilitate timely interventions, thus mitigating the risk of developing more severe conditions.

Approximately 25–35% of adults worldwide are affected by this health issue, with the prevalence increasing with age in both genders[4,5]. Recent studies have determined the pooled prevalence and incidence rates of MetS in the general population of Iran. The prevalence rate stands at 0.26 with a 95% Confidence Interval (CI) of 0.26 to 0.29, and the incidence rate is reported at 97.96 with a 95% CI ranging from 75.98 to 131.48[6]. Given its growing prevalence, MetS has become a crucial public health concern. Various studies have shown that individuals with MetS are at a higher risk of developing cardiovascular disease (CVD) and type 2 diabetes mellitus (T2DM), as well as having a higher mortality rate[7,8]. Therefore, early detection and intervention for MetS are critical in preventing the progression of such diseases.

Given the rapid advancements in technology, machine learning (ML) is increasingly recognized as a highly effective technology for data analysis. By leveraging ML and big data techniques, significant insights can be extracted from large datasets, thereby enhancing data-driven decision-making. This process is essential for effectively managing and making sense of large datasets, which aids in making well-informed, strategic decisions across various sectors[9]. In the medical field, ML offers significant insights by analyzing extensive clinical[10–12], imaging[13], and genomic data[14]. This analysis greatly improves the accuracy of diagnosing and classifying diseases, ultimately leading to innovative treatment methodologies.

In the field of ML for healthcare, traditional ML techniques such as linear regression (LR), decision trees (DT), and support vector machines (SVM) are extensively utilized and the application of these methods in predicting MetS is well-supported due to the substantial amount of research and data available[15]. A ML analysis using Random Forest (RF) and LR analysis on cohort data from Tehran identified a history of diabetes, body mass index (BMI), age, and female gender as key predictors of MetS. Their RF model achieved a high sensitivity of 0.97 and specificity of 0.99, indicating strong diagnostic performance for MetS prediction[16]. Additionally, a study using data from the Isfahan cohort employed SVM and DT algorithms, with the SVM model demonstrating superior performance. The SVM and DT algorithms employed in the study demonstrated robust classification

performance. Specifically, the SVM achieved a sensitivity of 0.774, specificity of 0.74, and overall accuracy of 0.757, while the DT method yielded slightly lower values, with a sensitivity of 0.758, specificity of 0.72, and accuracy of 0.739[17]. A recent review of ML methods in MetS prediction demonstrated that ML models achieved Area Under the Receiver Operating Characteristic Curve (AUROC) values ranging from 0.70 to 0.95 across various models and datasets. Traditional ML methods such as RF and LR have often been reported outperform deep learning techniques in terms of clinical interpretability. For instance, RF models achieved AUROC scores above 0.70 for MetS prediction using features such as waist circumference, blood glucose, and lipid profiles, while LR models attained AUROCs between 0.90 and 0.93 when using anthropometric features. In contrast, modern ML techniques like CNNs demonstrated higher accuracy (up to 93.3%) for classifying obesity, T2DM, and MetS but required larger sample sizes and increased computational complexity[18,19].

Liu et al. emphasized that traditional ML algorithms offer greater transparency and are better suited to structured clinical data, providing explicit feature importance for variables such as triglycerides and waist circumference. Conversely, contemporary deep learning models automate feature extraction and perform well on complex, multi-modal datasets but are often criticized as "black-box" systems due to limited interpretability. To address this, explainability frameworks (e.g., SHAP) are increasingly recommended. Overall, traditional ML approaches remain advantageous for clinical contexts requiring interpretability, while modern ML offers scalability and robustness at the expense of transparency and computational demands[19,20].

While many studies focus on anthropometric, lifestyle, and biochemical predictors in diagnosing MetS, the significance of liver function tests and hs-CRP has not been fully investigated. Notably, serum alanine aminotransferase (ALT) and aspartate aminotransferase (AST) are liver enzymes that increase in response to hepatocyte damage[21]. ALT is a more specific indicator of hepatocellular damage than AST. Elevated levels of alanine aminotransferase (ALT) are associated with an increased risk of developing CVD[22], obesity[23], insulin resistance[24], and MetS[25]. Although the underlying mechanisms are not yet fully understood, ALT-related non-alcoholic fatty liver disease (NAFLD) plays a significant role, particularly among men[26,27]. NAFLD is frequently linked to central obesity, type 2 diabetes, and dyslipidemia, and is therefore recognized as a hepatic manifestation of MetS[28,29].

The selection of liver function tests and hs-CRP as key biomarkers in this study is grounded in their well-established pathological association with MetS. Hepatic dysfunction, reflected in elevated liver enzymes such as ALT and AST, is closely linked to insulin resistance and NAFLD, both of which are central to MetS pathogenesis[30,31]. Similarly, hs-CRP, a marker of systemic inflammation, plays a critical role in MetS by mediating endothelial dysfunction and promoting atherogenesis[32]. Evidence suggests that chronic low-grade inflammation and hepatic steatosis contribute synergistically to metabolic dysregulation, further exacerbating MetS components such as dyslipidemia and hyperglycemia[33].By incorporating these biomarkers, this study aims to elucidate their diagnostic and prognostic relevance in MetS, thereby providing a more comprehensive understanding of the underlying metabolic and inflammatory disturbances.

Moreover, bilirubin, a byproduct of heme breakdown[34], has been inversely associated to dyslipidemia, hyperglycemia, and obesity[35]. Lower bilirubin levels have been associated with decreased insulin sensitivity and a higher occurrence of diabetes mellitus[36]. Bilirubin has been associated with a protective role that parallels the function of heme oxygenase-1 (HO-1) in mitigating insulin resistance[37]. Studies have demonstrated that HO-1 induction enhances insulin signaling pathways in various mouse models, including Zucker rats, leptin-deficient mice, leptin receptor-deficient mice, and diet-induced obese (DIO) mice[36,38–40]. Similarly, bilirubin administration has been shown to improve insulin sensitivity in obese mouse models[37].

Recent studies highlight the potential role of inflammation in triggering MetS[41]. hs-CRP, an acute-phase reactant synthesized by hepatocytes, increases in response to pro-inflammatory cytokines such as tumor necrosis factor-alpha (TNF-α) and interleukin-6 (IL-6)[42]. These cytokines have been found to disturb both lipid metabolism[43] and glucometabolism[44], thereby contributing to the pathogenesis of MetS. Moreover, CRP influences the insulin-induced phosphorylation of insulin receptor substrate 1 in laboratory settings, indicating that human CRP may interfere with glucose metabolism through insulin signaling pathways that control cellular glucose uptake[45]. Therefore, elevated hs-CRP levels are positively associated with a higher risk of T2DM, MetS, HTN, obesity, and CVD[41,46–51]. Kim et al. have identified serum hs-CRP as a significant chemical biomarker in predicting MetS using various ML methods[41].

Although previous studies have applied machine learning techniques for MetS prediction, they primarily focus on anthropometric and lifestyle factors, while biochemical markers such as liver function tests and hs-CRP remain underutilized. Given the well-established relationship between liver dysfunction, systemic inflammation, and MetS, our study integrates these biomarkers into a comprehensive predictive framework. Unlike previous models that rely solely on traditional ML approaches, we incorporate ensemble learning methods and deep learning models, such as CNNs, to improve prediction accuracy and robustness. This approach aims to bridge the gap in existing MetS prediction models by leveraging advanced ML techniques with clinically relevant biochemical markers.

To our knowledge, the combined effects of serum liver function tests—specifically ALT, AST, direct bilirubin (BIL.D), and total bilirubin (BIL.T)—combined with hs-CRP in predicting MetS through ensemble machine learning approaches have not been previously investigated. Therefore, this study aims to assess the influence of these biochemical markers on MetS, exploring their potential as novel predictors of MetS. Consequently, we developed an automated machine learning framework for predicting metabolic syndrome, leveraging serum liver function tests and high-sensitivity C-reactive protein (hs-CRP). This framework was developed using data from the Mashhad Stroke and Heart Atherosclerotic Disorder (MASHAD) cohort study, spanning a decade from 2010 to 2020[52,53]. It integrates both traditional and advanced machine learning models to enhance predictive accuracy. The initial stage focused on preprocessing and feature engineering techniques. Comprehensive preprocessing steps guaranteed the data's accuracy and reliability, while advanced feature extraction methods

enhanced the predictive capability of the models. During the second phase, we applied various ML models in our framework. This phase was critical for identifying suitable algorithms that could effectively diagnose MetS based on the extracted features. At the end, we performed a detailed assessment of the developed models to determine the most effective one for our purposes. This evaluation relied on a series of predefined metrics designed to measure both the accuracy and efficiency of each model. Existing ML-based MetS prediction models rely primarily on anthropometric, lifestyle, and basic biochemical markers such as fasting glucose and lipid profiles. However, emerging research highlights the role of liver function abnormalities and systemic inflammation in MetS progression. Despite this, most previous studies have overlooked the predictive power of serum liver enzymes and hs-CRP. This study fills this gap by integrating these biomarkers with ML techniques, leveraging their potential for more accurate and earlier MetS detection.

Methodology

This research developed an automated ML framework designed to predict metabolic syndrome, utilizing serum liver function tests (ALT, AST, BIL.D, and BIL.T), hs-CRP, along with demographic variables such as age and sex. The framework integrated a blend of traditional ML techniques—LR, DT, and SVM—and ensemble models like RF, BG, GB, and CNNs. This combination is designed to maximize prediction accuracy, as each model was specifically engineered to enhance the precision and effectiveness of predictions across various aspects of the MASHAD cohort dataset, which is described in Subsection 3.1. The framework developed in this research comprises two primary phases: data preprocessing and analysis, and model development. Initially, the first step involves data processing and preparation, which includes data cleaning, statistical analysis, feature engineering, and resampling, to guarantee its appropriateness for predictive modeling. Subsequently, the second phase focuses on the development and implementation of various ML models within the framework. This phase details the specific roles and interactions of each model. The final component, model validation, evaluates the constructed models using a range of techniques to determine their effectiveness and accuracy. The predictive models were developed and evaluated using Python 3.9, supported by robust libraries such as Scikit-learn, Pandas, and NumPy to ensure advanced analytical capabilities. An illustrative overview of the proposed framework is depicted in Fig. 1, offering visual clarity on the processes involved.

Data preprocessing

The efficacy of ML models in predicting MetS is significantly influenced by the quality and structure of the input data, which, in this study, included serum liver function tests and hs-CRP levels. In this step, to effectively ready

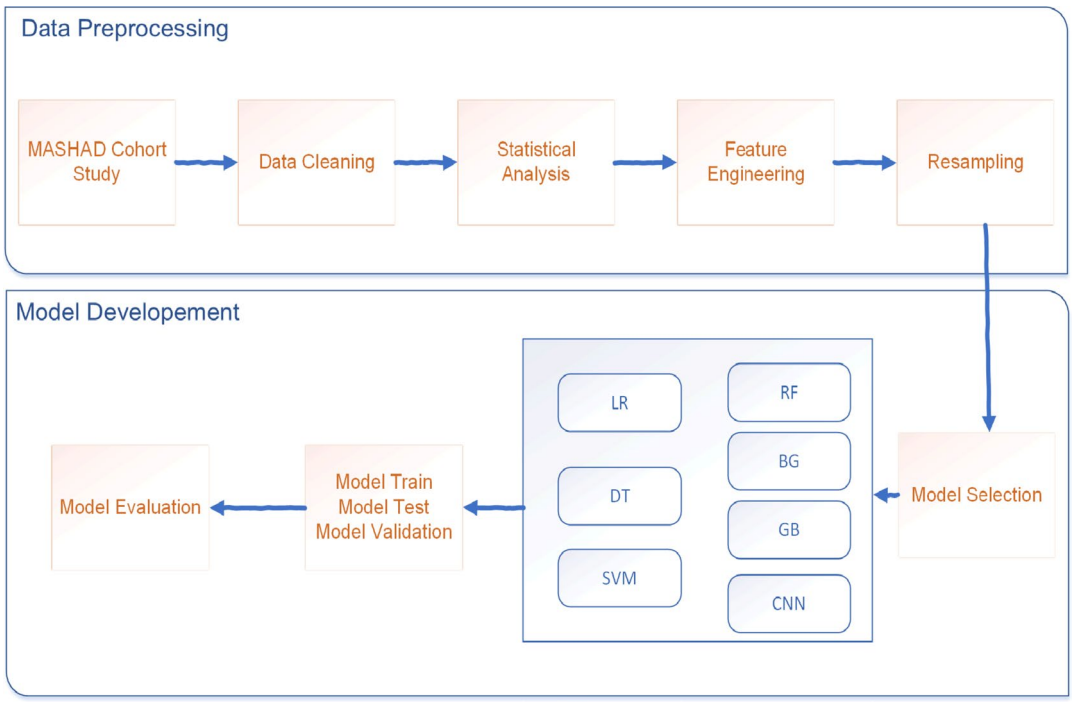

Fig. 1. Graphical representation of the proposed automated machine learning (ML) framework designed to predict Metabolic Syndrome (MetS) utilizing serum liver function tests (ALT, AST, BIL.D, BIL.T) and high-sensitivity C-reactive protein (hs-CRP). To enhance the precision of the analyses, seven diverse ML models were implemented: Linear Regression (LR), Decision Trees (DT), Support Vector Machines (SVM), Random Forests (RF), balanced bagging (BG), Gradient Boosting (GB), and Convolutional Neural Networks (CNNs). These models were systematically trained and evaluated on a cohort dataset encompassing 8,972 records with seven predictive variables. The dataset included 3,442 patients diagnosed with MetS and 5,530 healthy individuals without Metabolic Syndrome.

the input data for our ML-based framework, various preprocessing techniques were systematically structured into four distinct subsections. These steps include the detailed cleaning of data to remove inaccuracies, the transformation of features to suitable scales, and the handling of missing values and outliers to maintain the robustness of the model. The following subsections provide an in-depth account of these techniques, emphasizing their importance and their implementation to enhance the overall performance of the predictive models.

Data cleaning and statistical analysis

Data cleaning constituted the initial phase of the preprocessing process, characterized by the identification and elimination of incorrect data entries. This step involved correcting any inaccuracies or missing values in the dataset, as well as removing any values that appeared implausible and were likely the result of data entry errors. This approach ensures the dataset's integrity and reliability for subsequent analyses. Moreover, to optimize the performance of ML models, appropriate statistical analysis was utilized to mitigate discrepancies in feature values.

During data preprocessing, both Min-Max and Z-score normalization methods were applied and compared empirically across model types. Z-score normalization standardized the features by deducting the mean and dividing by the standard deviation, transforming the data to exhibit a zero mean and unit variance. Concurrently, Min-Max normalization adjusted the features to a consistent range from 0 to 1, normalizing the span of the features while maintaining the relative proportions of the original data values. Although Min-Max normalization aided convergence in deep learning models, our subsequent analysis in Sect. 3.3 showed that omitting normalization altogether yielded better predictive performance in Gradient Boosting models, particularly in terms of AUC, F1-score, and sensitivity. Therefore, the final framework incorporated 'no normalization' for optimal performance in ensemble models.

Moreover, detailed statistical analyses and a comprehensive description of the cohort dataset are provided in Subsection 3.1.

Feature engineering

Feature engineering and feature selection streamline the modeling process, reduce computational costs, enhance model interpretability, and can significantly boost the performance of ML models. Feature engineering is the process of transforming raw data into features that better represent the underlying problem to the predictive models, thereby increasing their accuracy on unseen data. Concurrently, feature selection plays a critical role by identifying the most relevant variables to use in model building. This reduces model complexity, while improving the model's performance and interpretability. In this study, Pearson correlation was initially used to identify linear associations between clinical variables and MetS. To further explore potential non-linear monotonic dependencies, we additionally applied SHAP (Shapley Additive Explanations) approach[54] to quantify both linear and non-linear feature contributions and interactions. SHAP is a powerful interpretability tool that helps explain the influence of individual variables on machine learning model predictions. This combined strategy provided deeper insights into the roles of weakly correlated variables—such as AST and BIL.T—by identifying their limited influence on model outputs beyond simple linear correlations.

SHAP values play a critical role in elucidating the contribution of each feature to the predictions made by machine learning models. This approach not only improves model accuracy but also enhances interpretability by precisely quantifying the impact of each feature.

As illustrated in Fig. 2, SHAP analysis has successfully identified significant predictors (age, hs-CRP, BIL.D, and sex) for MetS. The summary plot visually represents the importance of each feature, with features ranked by their SHAP values. The colors red and blue on the plot denote high and low impacts respectively. Adjacent to each feature's name, the mean SHAP value is displayed, emphasizing its relative significance in the model. Among the examined features, hs-CRP, BIL.D, ALT, and sex had the highest SHAP values and were retained in the final model. Features with mean absolute SHAP values below an empirical threshold of 0.05 (e.g., AST and BIL.T) were excluded from the final model to prevent overfitting.

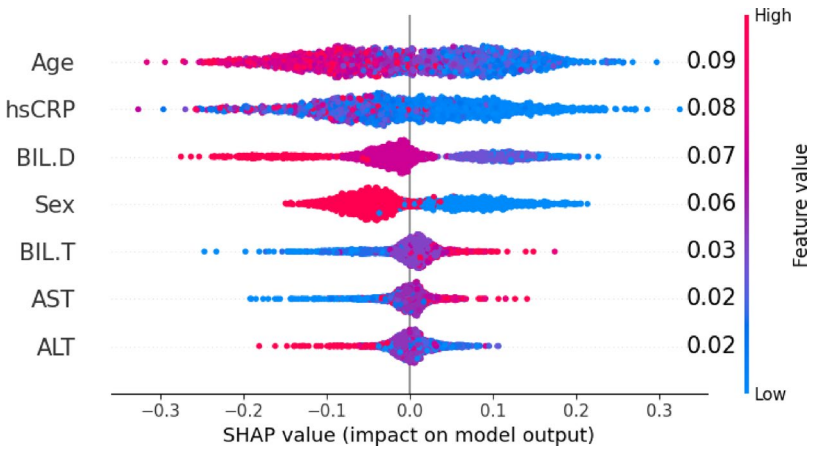

Fig. 2. Evaluation of features' importance by Shapley Additive Explanations (SHAP) summary plot.

Resampling imbalanced data

Resampling is a fundamental technique used to address class imbalance in datasets, which is a common issue in ML where the number of instances of one class significantly outweighs the other(s). This imbalance can lead to poor model performance, particularly in predicting the minority class. To address the challenges posed by imbalanced data within the MASHAD cohort study, this phase of the research incorporated a range of data resampling techniques. These techniques encompassed Random Over-Sampling (ROS), Random Under-Sampling (RUS), Synthetic Minority Oversampling Technique (SMOTE), Adaptive Synthetic Sampling (ADASYN), and Support Vector Machine Synthetic Minority Over-Sampling Technique (SVM-SMOTE)[55]. Each method aims to equalize the class distribution within the data, thus enhancing the robustness and predictive accuracy of the ML models. Among these approaches, SVM-SMOTE emerged as the most effective, exhibiting superior performance in handling data imbalances. SVM-SMOTE is a sophisticated technique that builds on the foundation of SMOTE by integrating the power of SVM classifiers to address class imbalance more effectively, particularly in complex classification scenarios. Consequently, SVM-SMOTE was chosen and applied to our training dataset to mitigate the imbalance issue. The impact of SVM-SMOTE on ML models is thoroughly discussed in Subsection 3.2.

ML models development

This section elaborates the ML models utilized in our automated framework. The framework incorporates seven well-known ML techniques such as LR, DT, and SVM with advanced sophisticated methods including RF, BG, GB, and CNNs to enhance predictive accuracy. Contrary to classical models (LR, DT, SVM), the meta-models such as ensemble ML models (RF, BG, GB, CNNs), achieve robust classifiers or regressors by constructing multiple weak models and integrating their outputs through voting or adaptively boosting schemes. These models attempt to generate precise predictions by optimally balancing the bias-variance trade-off and fostering generalized models. For more clarification, Table 1 provides further details on the classic and ensemble ML models utilized in our framework.

In the subsequent section, we describe the procedures employed to determine the optimal parameters and configurations for each model. Prior to delving into algorithmic analysis and performance evaluation, which will be extensively discussed in the following section, it is imperative to gain a thorough understanding of the individual models. This discussion is critical as it provides essential insights into how each model conforms and adapts to the data. To select optimal hyperparameters for ensemble models (RF, BB, GB), we employed grid search with Scikit-learn's GridSearchCV function. A predefined range of values was tested for each algorithm—such as n_estimators, max_depth, min_samples_split, learning_rate, and subsample—based on empirical practice and relevant literature. The final selected parameters (e.g., max_depth = 7 for RF and n_estimators = 300 for GB) were those that achieved the best validation performance in terms of F1-score and AUC. These optimized configurations were then used to evaluate model performance on the test set. Detailed configurations for each model are systematically presented in Table 2.

Models' performance evaluation

In this study, we assessed model performance using a range of metrics including sensitivity, specificity, F-measure, area under the curve (AUC), and error rate. These metrics, derived from the confusion matrix. The confusion

Algorithms	Description
Linear Regression (LR)	Linear regression is a fundamental statistical method designed to model the relationship between a dependent variable (the target of prediction) and one or more independent variables (known as predictors or features)[56]
Decision Tree (DT)	Decision tree models utilize a structured tree-like graph to depict decisions and their potential outcomes, simplifying the process for exploratory data analysis due to their clarity and ease of interpretation. These models are applicable for both classification and regression tasks. Structurally, a decision tree is comprised of a root node, numerous internal nodes, and several terminal nodes. The root and internal nodes, known collectively as nonterminal nodes, are interconnected in successive decision stages, while the terminal nodes signify the ultimate classifications[57]
Support Vector Machine (SVM)	Support vector machines are a robust supervised machine learning algorithm designed for binary classification. SVMs are engineered to identify the optimal boundary, or hyperplane, that effectively delineates various data classes. This is achieved through the kernel trick which is a sophisticated method that allows SVMs to operate with non-linearly separable data by transforming it into a higher-dimensional space where the separation can occur linearly[58].
Random Forest (RF)	Random forests is an ensemble learning method that constructs multiple decision trees, integrating their outputs to enhance prediction accuracy and stability. This technique employs internal metrics to monitor error, strength, and correlation, allowing for adjustments in response to increases in the number of features used for splitting[59]
Balanced Bagging (BG)	The Balanced bagging classifier is an ensemble method engineered to tackle class imbalance in classification tasks. It merges the principles of bagging, or bootstrap aggregating, with under-sampling the majority class, thereby ensuring each base classifier is trained on a balanced dataset. Bagging itself involves generating several subsets of the original data by sampling with replacement, and each subset is then used to train an independent base classifier[60,61]
Gradient Boosting (GB)	A gradient boosting is a powerful machine learning technique that builds predictive models in the form of an ensemble of weak prediction models, typically decision trees. It belongs to the family of boosting algorithms, which convert weak learners into strong ones in a sequential manner. This process involves adjusting each new model to better fit the residual errors of the combined existing models, thus gradually improving accuracy by refining a loss function[62].
Convolutional Neural Network (CNNs)	A neural network is a computational model inspired by the structure and function of the human brain[63]. It's composed of interconnected units called neurons, which process information. These neurons are organized in layers: • Input layer: Receives data. • Hidden layers: Process information through complex calculations. • Output layer: Produces the final result. Neural networks learn by adjusting the connections between neurons through a process called training. A Convolutional Neural Networks (CNNs) is a type of deep neural network that is especially powerful for processing data with a grid-like topology. CNNs are widely used in applications where the ability to automatically and adaptively learn spatial hierarchies of features is beneficial[64].

Table 1. ML models utilized in the proposed framework to predict MetS.

Methods	Hyperparameters	Inputs
Random Forest	random_state = 42, max_depth = 7, min_samples_split = 11, n_estimators = 75	hs-CRP, BIL.D Age, Sex
Balanced Bagging	estimator = RandomForestClassifier (random_state = 42, max_depth = 7, min_samples_split = 11, n_estimators = 75), n_estimators = 125, random_state = 42	hs-CRP, BIL.D Age, Sex
Gradient Boosting	n_estimators = 300, learning_rate = 0.05	hs-CRP, BIL.D Age, Sex
Convolutional Neural Network	models.Sequential([layers.Input(shape=(num_features,)), layers.Dense(64,activation='relu'), layers.Dense(32, activation='relu'), layers.Dense(num_classes, activation='sigmoid')]), optimizer='adam', loss='binary_crossentropy', epochs = 10, batch_size = 20, validation_split = 0.2	hs-CRP, BIL.D Age, Sex

Table 2. Parameters of ensemble models utilized in our Framework.

matrix categorizes predictions into four distinct outcomes: true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN), providing a clear depiction of both successes and errors made by the model. The derived metrics, each reflecting a different aspect of model performance, are calculated as follows:

$$Sensitivity\ (\ Recall\ or\ True\ Positive\ Rate) = TP/(TP+FN)$$

$$Specificity\ (True\ Negative\ Rate) = TN/(TN+FP)$$

$$Precison = TP/(TP+FP)$$

$$\mathrm{F-Measure\ (F1\ Score)} = 2 \times (\mathrm{Precision} \times \mathrm{Sensitivity}) / (\mathrm{Precision} + \mathrm{Sensitivity})$$

$$Error\ Rate = FP+FN/(TP+TN+FP+FN)$$

These metrics collectively assess the model's capability to accurately distinguish between individuals with and without MetS, where the AUC, derived from the receiver operating characteristic (ROC) curve, provides an aggregate measure of performance across all possible classification thresholds. This comprehensive evaluation aids in assessing the model's accuracy, its suitability for clinical application, and areas requiring improvement, which are further elaborated in the discussion section.

Experimental results

In this study, experiments were conducted using Python 3.9.1 within the Anaconda environment, employing libraries such as Scikit-learn, TensorFlow, Pandas, Keras, and NumPy. The computational analyses were performed on a MacBook Pro, which is equipped with a 2.2 GHz 6-Core Intel Core i7 Processor and 16 GB of RAM. Subsequent sections will detail the demographic characteristics of the study population, present the findings of the statistical analyses, and discuss their implications for the predictive models. Additionally, this section will evaluate and compare the performance of various models implemented within the framework.

Study population characterization

This subsection provides a comprehensive overview of the cohort dataset used in the analysis. The data originated from the MASHAD cohort study, which enrolled individuals from Mashhad, northeastern Iran, between 2010 and 2020[52]. The MASHAD cohort enrolled over 9,000 participants aged 35 to 65 years. Throughout the 6-year follow-up, cardiovascular events were identified via physical examinations and interviews conducted by two cardiologists. At the baseline, the study sample consisted of 5,530 healthy individuals and 3,442 subjects diagnosed with MetS, all of whom provided written informed consent. Comprehensive data collection included blood pressure, anthropometric, demographic, and other biochemical data were determined using BT3000 biochemical auto-analyzer. The ethical approval for this study was granted by the Mashhad University of Medical Sciences Ethical Committee (IR.MUMS.REC.85134). Table 3 summarizes the variables analyzed in this study.

Following the implementation of data processing and analytical techniques, a cohort of 8,972 participants was established for the final analysis, which included 3,442 individuals diagnosed with MetS and 5,530 without. As detailed in Table 3, there was a significant difference in gender distribution between the two groups ($p < 0.001$). Moreover, the mean ages of participants with and without MetS were 50.06 ± 7.97 and 46.86 ± 8.18, respectively, indicating a statistically significant difference ($p < 0.001$). Biochemical analysis revealed markedly elevated levels of hs-CRP, BIL.D, and ALT in those with MetS compared to their counterparts ($p < 0.001$). Conversely, BIL.T levels were significantly higher in individuals without MetS ($p < 0.001$). However, there was no significant difference in AST levels between the groups ($p = 0.550$).

Resampling results

This study examined the effect of diverse resampling techniques such as ROS, RUS, SMOTE, and ADASYN on the predictive performance of ML models in analytical framework. We conducted a comparative evaluation of multiple widely used resampling strategies for imbalanced classification. Each technique was applied to the training dataset, and model performance was assessed using standard evaluation metrics (AUC and sensitivity). As summarized in Table 4, SVM-SMOTE consistently outperformed other techniques based on the performance metrics. Notably, SVM-SMOTE yielded the highest AUC (0.72). While ADASYN also provided reasonable results (AUC = 0.70), its performance was marginally lower than SVM-SMOTE, particularly in sensitivity (0.75 vs. 0.73). The superior performance of SVM-SMOTE may be attributed to its ability to generate synthetic samples near the decision boundary by leveraging SVM hyperplanes, resulting in more effective representation

Variable	Total ($n=8972$)	Without MetS ($n=5530$)	With MetS ($n=3442$)	P-value*
Sex (Female)	5434 (60.60%)	3048 (56.10%)	2482 (70.20%)	< 0.001
Age	48.09 ± 8.25	46.86 ± 8.18	50.06 ± 7.97	< 0.001
ALT (U/L)	15.09 ± 5.24	14.89 ± 5.15	15.41 ± 5.36	< 0.001
AST (U/L)	21.91 ± 5.59	21.94 ± 5.59	21.86 ± 5.60	0.550
BIL.D (mg/dL)	0.27 ± 0.08	0.26 ± 0.08	0.30 ± 0.09	< 0.001
BIL.T (mg/dL)	0.42 ± 0.14	0.43 ± 0.14	0.41 ± 0.13	< 0.001
hs-CRP (mg/L)	2.93 ± 3.68	2.62 ± 3.49	3.44 ± 3.92	< 0.001

Table 3. Baseline characteristics of the Mashhad stroke and heart atherosclerotic disorder (MASHAD) Study[52]. ALT: Alanine Transaminase; AST: Aspartate Aminotransferase; BIL.D: Direct Bilirubin; BIL.T: Total Bilirubin; hs-CRP: High-sensitive C reactive protein. *P-value is computed based on t-tests for continuous data and Chi-square test for categorical data.

Resampling Method	AUC	Sensitivity	Specificity
Before Resampling	0.661	0.859	0.463
ROS	0.683	0.691	0.675
SMOTE	0.717	0.748	0.686
ADASYN	0.703	0.738	0.668
SVM-SMOTE	**0.72**	**0.75**	**0.69**
RUS	0.69	0.672	0.709

Table 4. Comparison of resampling methods based on AUC, sensitivity, and specificity in predicting MetS.

of the minority class in complex, high-dimensional feature spaces. Therefore, based on empirical evidence and theoretical justification, SVM-SMOTE was selected as the optimal resampling technique to address class imbalance in our predictive framework.

Moreover, this study examined the effect of resampling techniques on the predictive performance of ML models in analytical framework. As evidenced in the results (Fig. 3), the implementation of resampling method (SVM-SMOTE) notably influenced the AUC metrics across different models. The findings support that posits resampling as a beneficial strategy in managing class imbalance, particularly for ensemble and advanced ML models. The results also highlight the necessity of choosing appropriate resampling techniques lead to more accurate and reliable predictions of MetS, thus improving the diagnostic processes and subsequent management strategies for at-risk populations.

Statistical analysis results

To evaluate the effect of normalization techniques, we compared the performance of GB model trained with Z-score, Min-Max, and no normalization methods. As shown in Table 5, omitting normalization led to the highest F1-score (0.72) and sensitivity (0.74) on test data. These results indicate that no normalization is optimal for this use case.

In the context of ML, understanding the correlations between features is vital for effective feature selection and understanding the interdependencies among variables, which are key for refining models and enhancing their interpretability. In this study, we employed Pearson correlation analysis as a descriptive statistical tool to examine the relationships among serum liver function tests, hs-CRP, and MetS. This method calculates the Pearson correlation coefficient, which quantifies the strength and direction of a linear relationship between two continuous variables and varies between − 1 and + 1. Figure 4 presents the pairwise correlations between these variables, illustrating their linear interrelations. This step is essential for identifying significant correlations that may influence the predictive performance of the ML models utilized within our analytical framework.

Performance comparison of ML models

The ML framework developed for predicting MetS utilized a training dataset comprising 80% of the records. The models were subsequently evaluated using two distinct datasets: one constituting 10% of the records for testing and another 10% for validation purposes. Hyperparameter tuning was conducted on the validation set using grid search. Given the large cohort size (~ 9,000 samples), each 10% subset contained nearly 900 individuals, allowing for statistically stable performance estimates without needing k-fold cross-validation. This approach ensures unbiased evaluation, facilitates exact reproducibility, and aligns with standard practices in biomedical machine learning, particularly when comparing multiple models.

To address class imbalance within the training data, the SVM-SMOTE was implemented. The efficacy of the framework is critically assessed to predict outcomes for patients diagnosed with MetS compared to those without. The framework incorporates seven ML models: traditional models such as LR, DT, and SVM classifiers, and alongside ensemble methods including RF, BG, GB, and CNNs. For the first experiment, the performance of models for predicting MetS was compared using the metrics described in Sect. 2.3. The performance metrics are

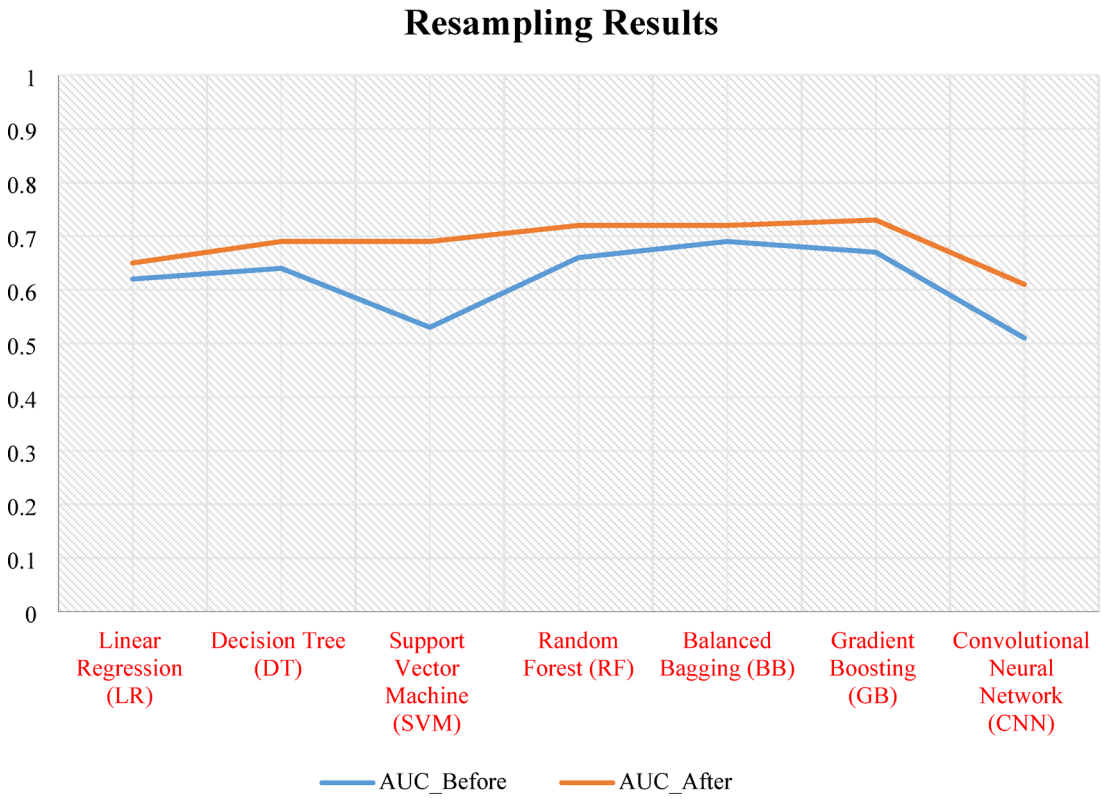

Fig. 3. AUC based on Performed Resampling.

Normalization	AUC	F1	Specificity	Sensitivity
None	0.718	0.726	0.691	0.746
Min-Max	0.695	0.690	0.714	0.677
Z-score	0.702	0.698	0.717	0.687

Table 5. Effect of normalization techniques on GB model Metrics.

F-measure, specificity, and sensitivity. Figure 5 indicates that the suggested ML models exhibited high values for the majority of measures. DT demonstrated superior results among the classic models in most metrics. Among ensemble models, both BG and GB were identified as superior compared to their counterparts. RF, BG and GB demonstrated superior results, achieving an F-measure of approximately 0.72. Additionally, the CNNs excelled in specificity, achieving a remarkable value of 83%.

In the second series of experiments, the robustness of developed framework was assessed by calculating the error rates of various ML models. The error rate, a crucial performance metric in ML, indicates the proportion of incorrect predictions made by a model relative to the total number of instances evaluated. This rate is calculated by dividing the number of incorrect predictions by the total predictions, and it is typically expressed as a percentage. This metric is vital for gauging a model's accuracy in correctly classifying or predicting outcomes. As illustrated in Fig. 6, classical ML models like DT and SVM demonstrated slightly lower error rates of 31%, compared to LR at 35%. Ensemble methods further reduced the error rate to approximately 27%. Notably, RF, which utilizes multiple DT, and BG, which counters class imbalance, significantly enhanced model robustness against overfitting and improved predictive accuracy. GB emerged as the most effective, with the lowest error rate of 27%, highlighting its capability in sequential model improvement by iteratively correcting errors from previous models, thus optimizing overall predictive performance.

Discussion

The present study aimed to evaluate the association between serum liver function tests and hs-CRP in predicting MetS using an automated ML framework. This approach combines sophisticated ensemble algorithms with conventional techniques, such as LR, DT, and SVM to significantly improve predictive accuracy. The proposed framework facilitates a comprehensive assessment of clinical biomarkers and their respective contributions to the progression of MetS. While previous studies have utilized machine learning methods to diagnose MetS, their focus has predominantly been on anthropometric and lifestyle factors. In contrast, the combined role of key

Fig. 4. Visualizing the heatmap. The figure shows the Pearson correlation coefficients between features and MetS. According to the pairwise correlation, ALT and AST exhibit a significant positive correlation with a coefficient of 0.37, Similarly, BIL.T and BIL.D are positively correlated with a coefficient of 0.24. In the context of MetS, there is a significant positive correlation between BIL.D and MetS at 0.20, and between age and MetS at 0.19. On the other hand, MetS shows slight negative correlations with AST and BIL.T, indicating weak inverse relationships in these cases.

biochemical markers, such as liver function tests and hs-CRP, remains underexplored in this context. Given the established associations between liver dysfunction, systemic inflammation, and MetS, this study incorporates these biomarkers into a sophisticated and automated predictive model, offering a more nuanced understanding of their interplay in MetS development.

ML has become pivotal in identifying risk factors associated with MetS by analyzing lifestyle choices, anthropometric data, and blood markers. A study conducted in Tehran utilizing RF and LR identified a history of diabetes, elevated BMI, age, and low monounsaturated fat intake as significant predictors of MetS. However, this study had a relatively small sample size of 3,048 participants and a follow-up period of only three years, compared to the ten years in our study. Unlike the Tehran study, which only utilized RF and LR, our research employs a combination of both novel and traditional ML methods, enhancing our understanding of MetS prediction within an Iranian population. To further refine the accuracy of our model, we implemented resampling techniques and conducted performance comparisons among the employed methods[16].

In another machine learning study using cohort data from Isfahan, Karimi et al. found that triglyceride levels had the highest predictive value when analyzed using a DT algorithm. They also noted that the accuracy of predictions significantly improved when triglyceride levels were combined with BMI. In the study, the DT and SVM methods achieved accuracies of 0.75 and 0.76, respectively[17]. Notably, this investigation involved a smaller cohort of 2,107 participants and considered a total of 22 variables, yet it did not include serum liver factors or high-sensitivity C-reactive protein (hs-CRP). Moreover, it was limited to using only two machine learning methods—DT and SVM—compared to the broader techniques employed in our study. This places our study among the pioneering efforts to assess Mets prediction using a diverse machine learning technique, implemented within a large-scale cohort comprising over 9,000 participants.

Ensemble learning methods in our study, particularly GB and BG, exhibited superior specificity and sensitivity compared to traditional ML methods. Our integrative framework combines both classical and modern algorithms, leveraging the strengths of each model. Traditional ML methods such as DT and SVM demonstrated a comparatively higher error rate of 31%, whereas ensemble methods reduced the error rate to approximately 27%. Notably, RF, leveraging multiple DT, and BG, aimed at correcting class imbalances, significantly enhanced model resilience against overfitting and boosted predictive accuracy. Among these, GB was the most effective one and achieving the lowest error rate. Gradient Boosting (GB) demonstrated superior performance through iterative model refinement and optimization of residual errors. Additionally, CNNs demonstrated superiority in specificity, achieving a remarkable rate of 83%. In our SHAP analysis, age, hs-CRP, BIL.D, and sex significantly

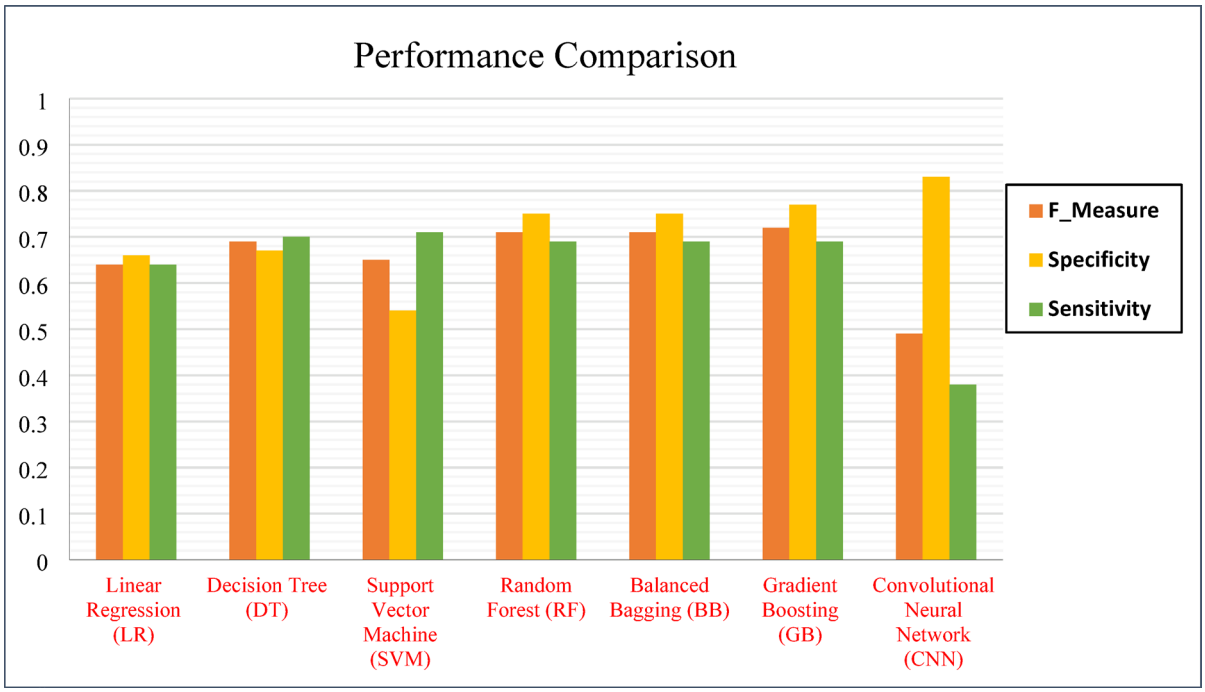

Fig. 5. Evaluating the ML Performance. The ML models including linear regression, decision tree, support vector machine, random forests, balanced bagging, gradient boosting, and convolutional neural networks compared using the three metrices (F-measure, specificity, and sensitivity). Each model's performance is represented by bars in different colors for each metric. The F-measure is depicted in yellow, specificity in orange, and sensitivity in green. Random Forest (RF), Balanced Bagging (BG), Gradient Boosting (GB), and Convolutional Neural Network (CNN) demonstrated superior results among the developed ML models based on the specificity.

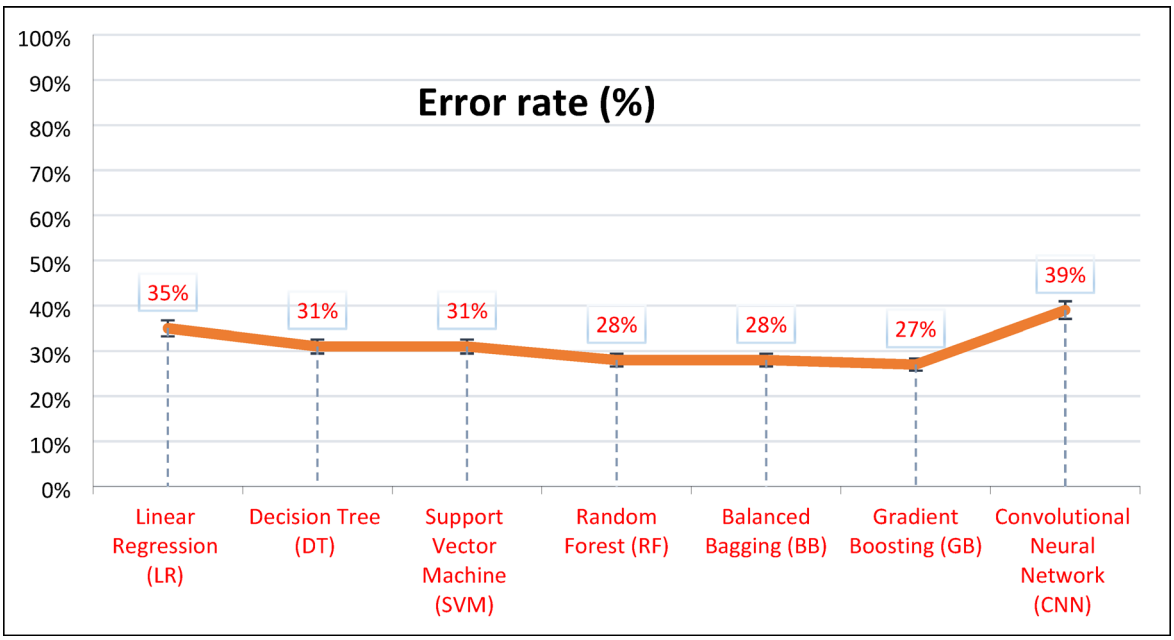

Fig. 6. Comparison of Error Rates (expressed as percentages) for ML Methods utilized in our Framework.

influenced the predictions of our ML models. Notably, hs-CRP emerged as a pivotal biomarker for predicting MetS within our specialized ML framework. Based on the pairwise analysis detailed in Subsection 3.3, the study revealed positive correlations between MetS and BIL.D, age, and sex. Conversely, a slight negative correlation was observed between MetS and BIL.T. However, these correlations were relatively modest and therefore require additional investigation to assess their clinical relevance.

Based on the existing literature, Bilirubin may play a significant protective role against MetS due to its multifaceted biological properties, including its potent antioxidant[65], anti-inflammatory effects[66], and insulin-sensitizing effects[36]. Another investigation highlights that reduced bilirubin levels are associated with several metabolic disturbances, including elevated blood pressure[67], increased glucose levels[68], and impaired lipid metabolism[69] all of which are key contributors to the development of MetS. Studies often focus on total serum bilirubin, which encompasses both direct and indirect bilirubin. In addition, bilirubin levels have been found to be inversely correlated with adiposity in obese individuals, while its catabolic byproduct, urobilin, has been positively linked to insulin resistance[70].

Our findings in this study suggest a negative association between BIL.T and MetS, contrasted by a positive correlation with BIL.D and MetS. Our findings align with previous studies indicating an inverse relationship between total bilirubin and MetS. Like the Korean cohort study by Lee et al. (2014)[71], we observed that lower BIL.T levels were associated with higher MetS risk. Moreover, a Japanese study by Oda and Aizawa (2013) also reported a negative link between BIL.T and MetS but clarified that low BIL.T levels were not independently predictive of increased MetS risk[72]. On the other hand, Li et al. (2017)[73] found a negative association between BIL.D and MetS, while our study observed a positive correlation.

This discrepancy may arise from the fact that bilirubin levels—particularly BIL.D—can increase in response to hepatic stress associated with MetS. As MetS progresses to NAFLD, bilirubin conjugation may initially increase; however, in later stages, hepatic dysfunction and cholestasis can occur, elevating direct bilirubin levels. Additionally, obesity and insulin resistance may lead to mild bile duct obstruction—particularly due to hepatic steatosis—which can elevate BIL.D levels independently of total bilirubin concentrations[74]. In parallel, ALT, a well-established marker of hepatic dysfunction, was also found to be positively associated with MetS in our study. Moreover, the elevated risk of CVD among individuals with MetS has been consistently documented[8]. Cardiovascular conditions such as heart failure may impair liver function through several hepatological mechanisms. Specifically, hyperbilirubinemia in the context of congestive heart failure is commonly attributed to hepatocellular dysfunction, passive hepatic congestion, bile thrombi formation, hemolysis, and the use of certain medications[75].

Another possible explanation lies in population-specific biological factors. Variability in bilirubin metabolism, such as differences in hepatic enzyme function and bilirubin conjugation capacity, may result in differing circulating BIL.D levels across populations. Genetic polymorphisms—particularly those involving the UGT1A1 gene—can influence bilirubin clearance and exhibit substantial ethnic variation, potentially leading to cohort-specific metabolic responses[76]. Moreover, environmental influences, including dietary patterns, lifestyle behaviors, and the prevalence of comorbidities such as obesity or diabetes, may further modulate the role of bilirubin in metabolic regulation. These factors collectively underscore the need for further investigation into the mechanisms linking direct bilirubin (BIL.D) metabolism with MetS development across diverse ethnic populations[76,77].

hs-CRP is widely recognized biomarker for inflammation, often found elevated in conditions such as diabetes, CVD, and MetS. Notably, a Korean cohort study highlighted hs-CRP's significant role in the development of MetS among women, regardless of body fat levels. This gender-specific impact might stem from differences in sex hormones and fat distribution, affecting inflammatory responses[78]. Several mechanisms have been suggested to determine the role of hs-CRP in MetS development. hs-CRP is produced in response to pro-inflammatory cytokines, which can lead to disruptions in lipid metabolism[73,79]. Additionally, hs-CRP has been associated with insulin resistance[80] and is thought to interfere with insulin-regulated metabolic pathways[81]. ALT and AST are critical liver enzymes routinely assessed in liver function tests and recognized as key biomarkers of liver damage[82]. ALT is considered a more specific indicator of liver functionality and is commonly used as an indirect marker for liver inflammation or injury. In contrast, AST has lower specificity because it is also present in other tissues, including the heart and muscles, which can complicate the interpretation of liver function results[83]. Non-alcoholic fatty liver disease (NAFLD) is a plausible cause of abnormal liver enzyme levels and can lead to asymptomatic elevations of ALT levels[27].

NAFLD has been links to several metabolic disorders, including obesity, hyperglycemia, hypertension, and dyslipidemia, all of which are components of MetS. This association supports the concept that NAFLD represents a hepatic manifestation of MetS[28]. Numerous cross-sectional studies have highlighted the relationship between elevated ALT and AST levels and MetS[26]. Consequently, increased ALT and AST levels may be considered potential risk factors for the development of MetS[28]. Our study population exhibited higher ALT levels in patients with MetS; however, the AST levels did not significantly differ among the studied groups. Additionally, our Pearson correlation analysis revealed a stronger correlation between ALT and MetS at 0.048, compared to AST, which was −0.0063. A related ML study employing LR in a Chinese population demonstrated that individuals with the highest ALT and AST group were predominantly male, often current smokers or drinkers, and exhibited higher levels of blood pressure, waist circumference (WC), BMI, total cholesterol, triglycerides, LDL-C, and fasting plasma glucose. However, the study noted that within the normal reference range, AST levels were not associated with MetS[84].

Another methodological machine learning study, which employed LR, ANN, SVM, and Classification and Regression Tree (CART), identified ALT as a significant predictive biomarker in a model evaluating the four-year risk of MetS in adults. Interestingly, the logistic regression model outperformed the other models, including CART and SVM, and demonstrated slightly better accuracy than the ANN model. The authors suggested that

this superior performance might be due to the predominance of independent, linear effects among the most influential prognostic factors, which logistic regression is particularly well-suited to capture[85].This finding highlights the importance of selecting appropriate modeling techniques based on the underlying structure and relationships within the data.

In the current study, we integrated classical and advanced machine learning techniques, combined with robust preprocessing methods, to enhance predictive accuracy and effectively address dataset imbalances. The performance of the models was evaluated using a comprehensive set of well-known metrics, including sensitivity, specificity, F-measure, area under the curve (AUC), and error rates. A key strength of this research lies in its emphasis on understudied biomarkers, which have not been extensively explored in prior studies, thereby contributing novel insights to the field, and advancing the understanding of their predictive potential. On the other hand, this study has several limitations that warrant careful consideration. Its cross-sectional design inherently restricts the ability to establish causal relationships between liver function tests, hs-CRP, and MetS. While the model achieved moderate accuracy (approximately 75%), it offered valuable insights into the associations between MetS and less frequently studied biomarkers, such as BIL.D and BIL.T. Furthermore, the research was conducted within an Iranian cohort, which may limit the generalizability of the findings to other ethnic or geographic populations. Additionally, the gender distribution of the sample was uneven, with a higher proportion of females (59% female, 41% male), potentially introducing a bias toward female-related outcomes. These limitations highlight the need for further longitudinal and multi-ethnic studies to validate and expand upon these findings.

The high specificity of the proposed ML framework highlights its strong potential for practical clinical implementation. Integrating the model into electronic health records (EHRs) could facilitate automated MetS risk scoring during routine blood testing—particularly by utilizing standard liver function biomarkers and hs-CRP levels—to enable early identification of high-risk individuals. As a clinical decision-support tool, the framework could inform personalized interventions, whether lifestyle-based or pharmacological, and provide real-time risk assessment through an interactive clinician dashboard. Beyond primary care, embedding the model into public health initiatives, such as workplace wellness programs, would promote scalable, cost-effective screening and accelerate early detection. Broad adoption of this framework could contribute to reducing the societal burden of MetS, cardiovascular disease, and type 2 diabetes through targeted, preventative strategies.

Conclusion

This research was conducted using the MASHAD cohort, which consists of 9,704 participants over a 10-year follow-up period. This extensive sample size and follow-up duration provided a reliable foundation for the study, ensuring the data's robustness and comprehensiveness for detailed analysis.

In summary, this study introduced a robust ML framework designed to predict MetS using serum liver function tests and hs-CRP. By employing both classical and ensemble ML algorithms, we demonstrated the potential to significantly improve the predictive accuracy of models for identifying MetS. Key liver function markers, including ALT, AST, BIL.D, and BIL.T, along with hs-CRP, emerged as novel and effective biomarkers for assessing MetS risk. Notably, the study identified hs-CRP and BIL.D levels as significant biochemical predictors of MetS, underscoring their potential value in early screening and risk stratification strategies. These findings contribute to the growing body of evidence supporting the use of advanced computational approaches in enhancing the early detection and diagnose of metabolic disorders. In clinical practice, the proposed machine learning framework could be integrated into electronic health record (EHR) systems to automatically flag individuals at high risk for MetS based on standard liver function tests and hs-CRP levels. As a clinical decision-support tool, it can guide early interventions, helping clinicians tailor lifestyle or pharmacological strategies to individual biomarker profiles. Moreover, by facilitating scalable, low-cost, and non-invasive screening, the framework has the potential to be deployed in large-scale population health initiatives. Its adoption could play a pivotal role in reducing the societal burden of Mets, cardiovascular disease, and type 2 diabetes through earlier detection and prevention.

Data availability

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

Received: 4 March 2025; Accepted: 10 June 2025
Published online: 01 July 2025

References

1. Alberti, K. G. M. M., Zimmet, P. & Shaw, J. The metabolic syndrome—a new worldwide definition. *Lancet* **366**, 1059–1062. https://doi.org/10.1016/S0140-6736(05)67402-8 (2005).
2. Grundy, S. M. et al. Diagnosis and management of the metabolic syndrome: an American heart association/national heart, lung, and blood Institute scientific statement. *Circulation* **112**, 2735–2752. https://doi.org/10.1161/CIRCULATIONAHA.105.169404 (2005).
3. den Engelsen, C. et al. High-sensitivity C-reactive protein to detect metabolic syndrome in a centrally obese population: a cross-sectional analysis. *Cardiovasc. Diabetol.* **11**, 1–7. https://doi.org/10.1186/1475-2840-11-25 (2012).
4. Saklayen, M. G. The global epidemic of the metabolic syndrome. *Curr. Hypertens. Rep.* **20**, 1–8. https://doi.org/10.1007/s11906-018-0812-z (2018).
5. Farmanfarma, K. K. et al. Prevalence of metabolic syndrome in iran: A meta-analysis of 69 studies. *Diabetes Metabolic Syndrome: Clin. Res. Reviews.* **13**, 792–799. https://doi.org/10.1016/j.dsx.2018.11.055 (2019).
6. Fatahi, A., Doosti-Irani, A. & Cheraghi, Z. Prevalence and incidence of metabolic syndrome in iran: a systematic review and meta-analysis. *Int. J. Prev. Med.* **11**, 64. https://doi.org/10.4103/ijpvm.IJPVM_489_18 (2020).

7. Lu, J. et al. Metabolic syndrome among adults in china: the 2010 China noncommunicable disease surveillance. *J. Clin. Endocrinol. Metab.* **102**, 507–515. https://doi.org/10.1210/jc.2016-2477 (2017).
8. Kastorini, C. M. et al. Metabolic syndrome and 10-year cardiovascular disease incidence: the ATTICA study. *Nutr. Metabolism Cardiovasc. Dis.* **26**, 223–231. https://doi.org/10.1016/j.numecd.2015.12.010 (2016).
9. Wu, X., Zhu, X., Wu, G. Q. & Ding, W. Data mining with big data. *IEEE Trans. Knowl. Data Eng.* **26**, 97–107. https://doi.org/10.1109/TKDE.2013.109 (2013).
10. Ibrahim, M., Beneyto, A., Contreras, I. & Vehi, J. An ensemble machine learning approach for the detection of unannounced meals to enhance postprandial glucose control. *Comput. Biol. Med.* 108154. https://doi.org/10.1016/j.compbiomed.2024.108154 (2024).
11. Ahari, R. K. et al. Association of atherosclerosis indices, serum uric acid to high-density lipoprotein cholesterol ratio and triglycerides-glucose index with hypertension: A gender-disaggregated analysis. *J. Clin. Hypertens.* https://doi.org/10.1016/j.compbiomed.2024.108154 (2024).
12. Kolahi Ahari, R. et al. Association of three novel inflammatory markers: lymphocyte to HDL-C ratio, High-Sensitivity C-Reactive protein to HDL-C ratio and High-Sensitivity C-Reactive protein to lymphocyte ratio with metabolic syndrome. *Endocrinol. Diabetes Metab.* **7**, e00479. https://doi.org/10.1002/edm2.479 (2024).
13. Bloch, L., Friedrich, C. M. & Initiative, A. D. N. Systematic comparison of 3D deep learning and classical machine learning explanations for alzheimer's disease detection. *Comput. Biol. Med.* **170**, 108029. https://doi.org/10.1016/j.compbiomed.2024.108029 (2024).
14. Jablonka, K. M., Ongari, D., Moosavi, S. M. & Smit, B. Big-data science in porous materials: materials genomics and machine learning. *Chem. Rev.* **120**, 8066–8129. https://doi.org/10.1021/acs.chemrev.0c00004 (2020).
15. Kakudi, H. A., Loo, C. K. & Moy, F. M. Diagnosis of metabolic syndrome using machine learning, statistical and risk quantification techniques: A systematic literature review. *MedRxiv* **2020.06.01.20119339** https://doi.org/10.1101/2020.06.01.20119339 (2020).
16. Hosseini-Esfahani, F. et al. Using machine learning techniques to predict factors contributing to the incidence of metabolic syndrome in tehran: cohort study. *JMIR Public. Health Surveill.* **7**, e27304. https://doi.org/10.2196/27304 (2021). (accessed August 28, 2024).
17. Karimi-Alavijeh, F., Jalili, S. & Sadeghi, M. Predicting metabolic syndrome using decision tree and support vector machine methods. *ARYA Atheroscler.* **12**, 146 (2016). PMID: 27752272; PMCID: PMC5055373.
18. Eyvazlou, M. et al. Prediction of metabolic syndrome based on sleep and work-related risk factors using an artificial neural network. *BMC Endocr. Disord.* **20**, 1–11. https://doi.org/10.1186/s12902-020-00645-x (2020).
19. Liu, J. et al. Integrating artificial intelligence in the diagnosis and management of metabolic syndrome: A comprehensive review. *Diabetes Metab. Res. Rev.* **41**, e70039. https://doi.org/10.1002/dmrr.70039 (2025).
20. Churpek, M. M. et al. Multicenter comparison of machine learning methods and conventional regression for predicting clinical deterioration on the wards. *Crit. Care Med.* **44**, 368–374. https://doi.org/10.1097/CCM.0000000000001571 (2016).
21. Hanley, A. J. G. et al. Elevations in markers of liver injury and risk of type 2 diabetes: the insulin resistance atherosclerosis study. *Diabetes* **53**, 2623–2632. https://doi.org/10.2337/diabetes.53.10.2623 (2004).
22. Yokoyama, M. et al. Association of the aspartate aminotransferase to Alanine aminotransferase ratio with BNP level and cardiovascular mortality in the general population: the Yamagata study 10-year follow-up. *Dis. Markers.* **2016**, 4857917 (2016).
23. Klein, M. et al. Alanine transferase: an independent indicator of adiposity related comorbidity risk in youth: 丙氨酸转移酶: 一个年轻人肥胖相关合并症风险的独立指标. *J. Diabetes.* **7**, 649–656 (2015).
24. Hanley, A. J. G., Wagenknecht, L. E., Festa, A., D'Agostino, R. B. Jr & Haffner, S. M. Alanine aminotransferase and directly measured insulin sensitivity in a multiethnic cohort: the insulin resistance atherosclerosis study. *Diabetes Care.* **30**, 1819–1827 (2007).
25. Ballestri, S. et al. Nonalcoholic fatty liver disease is associated with an almost twofold increased risk of incident type 2 diabetes and metabolic syndrome. Evidence from a systematic review and meta-analysis. *J. Gastroenterol. Hepatol.* **31**, 936–944 (2016).
26. Bekkelund, S. I. Serum Alanine aminotransferase activity and risk factors for cardiovascular disease in a Caucasian population: the Tromsø study. *BMC Cardiovasc. Disord.* **21**, 1–7. https://doi.org/10.1186/s12872-020-01826-1 (2021).
27. Jalilian, M., Rasad, R. & Rotbeh, A. Fatty liver disease in overweight and obese Iranian children: comprehensive systematic review and meta-analysis. *Obes. Med.* 100455. https://doi.org/10.1016/j.obmed.2022.100455 (2022).
28. Rinaldi, L. et al. Mechanisms of non-alcoholic fatty liver disease in the metabolic syndrome. A narrative review. *Antioxidants* **10**, 270. https://doi.org/10.3390/antiox10020270 (2021).
29. Yki-Järvinen, H. Non-alcoholic fatty liver disease as a cause and a consequence of metabolic syndrome. *Lancet Diabetes Endocrinol.* **2**, 901–910. https://doi.org/10.1016/S2213-8587(14)70032-4 (2014).
30. Ghotbi, S. et al. Evaluation of elevated serum liver enzymes and metabolic syndrome in the PERSIAN Guilan cohort study population. *Heliyon* **10**(11), e32449 (2024).
31. Makri, E., Goulas, A. & Polyzos, S. A. Epidemiology, pathogenesis, diagnosis and emerging treatment of nonalcoholic fatty liver disease. *Arch. Med. Res.* **52**, 25–37 (2021).
32. Devaraj, S., Singh, U. & Jialal, I. Human C-reactive protein and the metabolic syndrome. *Curr. Opin. Lipidol.* **20**, 182–189 (2009).
33. Zheng, H., Sechi, L. A., Navarese, E. P., Casu, G. & Vidili, G. Metabolic dysfunction-associated steatotic liver disease and cardiovascular risk: a comprehensive review. *Cardiovasc. Diabetol.* **23**, 346 (2024).
34. Rawal, R. et al. A comprehensive review of bilirubin determination methods with special emphasis on biosensors. *Process Biochem.* **89**, 165–174. https://doi.org/10.1016/j.procbio.2019.10.034 (2020).
35. Nano, J. et al. Association of Circulating total bilirubin with the metabolic syndrome and type 2 diabetes: a systematic review and meta-analysis of observational evidence. *Diabetes Metab.* **42**, 389–397. https://doi.org/10.1016/j.diabet.2016.06.002 (2016).
36. Liu, J. et al. Bilirubin increases insulin sensitivity by regulating cholesterol metabolism, adipokines and PPARγ levels. *Sci. Rep.* **5**, 1–12. https://doi.org/10.1038/srep09886 (2015).
37. Dong, H. et al. Bilirubin increases insulin sensitivity in leptin-receptor deficient and diet-induced obese mice through suppression of ER stress and chronic inflammation. *Endocrinology* **155**, 818–828 (2014).
38. Li, M. et al. Interdiction of the diabetic state in NOD mice by sustained induction of Heme oxygenase: possible role of carbon monoxide and bilirubin. *Antioxid. Redox Signal.* **9**, 855–863 (2007).
39. Nicolai, A. et al. Heme oxygenase-1 induction remodels adipose tissue and improves insulin sensitivity in obesity-induced diabetic rats. *Hypertension* **53**, 508–515 (2009).
40. Shakeri-Manesch, S. et al. Diminished upregulation of visceral adipose Heme oxygenase-1 correlates with waist-to-hip ratio and insulin resistance. *Int. J. Obes.* **33**, 1257–1264 (2009).
41. Jeong, H. et al. C reactive protein level as a marker for dyslipidaemia, diabetes and metabolic syndrome: results from the Korea National health and nutrition examination survey. *BMJ Open.* **9**, e029861. https://doi.org/10.1136/bmjopen-2019-029861 (2019).
42. Xue, Q. et al. Association between baseline and changes in high-sensitive C-reactive protein and metabolic syndrome: a nationwide cohort study and meta-analysis. *Nutr. Metab. (Lond).* **19**, 1–12. https://doi.org/10.1186/s12986-021-00632-6 (2022).
43. Maury, E. & Brichard, S. M. Adipokine dysregulation, adipose tissue inflammation and metabolic syndrome. *Mol. Cell. Endocrinol.* **314**, 1–16 (2010).
44. McCracken, E., Monaghan, M. & Sreenivasan, S. Pathophysiology of the metabolic syndrome. *Clin. Dermatol.* **36**, 14–20 (2018).
45. D'Alessandris, C., Lauro, R., Presta, I. & Sesti, G. C-reactive protein induces phosphorylation of insulin receptor substrate-1 on Ser 307 and Ser 612 in L6 myocytes, thereby impairing the insulin signalling pathway that promotes glucose transport. *Diabetologia* **50**, 840–849 (2007).

46. Hong, G. et al. High-sensitivity C-reactive protein leads to increased incident metabolic syndrome in women but not in men: a five-year follow-up study in a Chinese population. *Diabetes Metab. Syndr. Obes.* **13**, 581. https://doi.org/10.2147/DMSO.S241774 (2020).
47. Gao, X. et al. C-reactive protein as a moderator and insulin resistance as a mediator for the association between Endothelin-1 and dysglycemia among African americans: Jackson heart study. *Obes. Med.* **33**, 100435. https://doi.org/10.1016/j.obmed.2022.100435 (2022).
48. Jayedi, A. et al. Inflammation markers and risk of developing hypertension: a meta-analysis of cohort studies. *Heart* **105**, 686–692. https://doi.org/10.1136/heartjnl-2018-314216 (2019).
49. Mogharnasi, M., TaheriChadorneshin, H. & Abbasi-Deloei, N. Effect of exercise training type on plasma levels of vaspin, nesfatin-1, and high-sensitivity C-reactive protein in overweight and obese women. *Obes. Med.* **13**, 34–38. https://doi.org/10.1016/j.obmed.2018.12.006 (2019).
50. Kolahi Ahari, R. et al. Serum uric acid to high-density lipoprotein ratio as a novel indicator of inflammation is correlated with the presence and severity of metabolic syndrome: A large-scale study. *Endocrinol. Diabetes Metab.* **6**, e446. https://doi.org/10.1002/edm2.446 (2023).
51. Saberi-Karimian, M. et al. Data mining approaches for type 2 diabetes mellitus prediction using anthropometric measurements. *J. Clin. Lab. Anal.* **37**, e24798. https://doi.org/10.1002/jcla.24798 (2023).
52. Ghayour-Mobarhan, M. et al. Mashhad stroke and heart atherosclerotic disorder (MASHAD) study: design, baseline characteristics and 10-year cardiovascular risk Estimation. *Int. J. Public. Health.* **60**, 561–572. https://doi.org/10.1007/s00038-015-0679-6 (2015).
53. Mansoori, A. et al. Prediction of type 2 diabetes mellitus using hematological factors based on machine learning approaches: a cohort study analysis. *Sci. Rep.* **13**, 663. https://doi.org/10.1038/s41598-022-27340-2 (2023).
54. Nohara, Y., Matsumoto, K., Soejima, H. & Nakashima, N. Explanation of machine learning models using Shapley additive explanation and application for real data in hospital. *Comput. Methods Programs Biomed.* **214**, 106584. https://doi.org/10.1016/j.cmpb.2021.106584 (2022).
55. Piri, S., Delen, D. & Liu, T. A synthetic informative minority over-sampling (SIMO) algorithm leveraging support vector machine to enhance learning from imbalanced datasets. *Decis. Support Syst.* **106**, 15–29. https://doi.org/10.1016/j.dss.2017.11.006 (2018).
56. Su, X., Yan, X. & Tsai, C. Linear regression. *Wiley Interdiscip Rev. Comput. Stat.* **4**, 275–294. https://doi.org/10.1002/wics.1198 (2012).
57. Swain, P. H. & Hauska, H. The decision tree classifier: design and potential. *IEEE Trans. Geoscience Electron.* **15**, 142–147. https://doi.org/10.1109/TGE.1977.6498972 (1977).
58. Cristianini, N. & Ricci, E. Support vector machines, in: Encyclopedia of Algorithms, Springer-, : 928–932. https://doi.org/10.1109/5254.708428. (2008).
59. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32. https://doi.org/10.1023/A:1010933404324 (2001).
60. Błaszczyński, J. & Stefanowski, J. Actively balanced bagging for imbalanced data, in: Foundations of Intelligent Systems: 23rd International Symposium, ISMIS 2017, Warsaw, Poland, June 26–29, Proceedings 23, Springer, 2017: pp. 271–281. (2017). https://doi.org/10.1007/978-3-319-60438-1_27
61. Błaszczyński, J. & Stefanowski, J. Improving bagging ensembles for class imbalanced data by active learning. *Adv. Feature Selection Data Pattern Recognit.* 25–52. https://doi.org/10.1007/978-3-319-67588-6_3 (2018).
62. Natekin, A. & Knoll, A. Gradient boosting machines, a tutorial. *Front. Neurorobot.* **7**, 21. https://doi.org/10.3389/fnbot.2013.00021 (2013).
63. McCulloch, W. S. & Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* **5**, 115–133. https://doi.org/10.1007/BF02478259 (1943).
64. Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. *Nature* **323**, 533–536. https://doi.org/10.1038/323533a0 (1986).
65. Lin, J. P. et al. Association between the UGT1A1* 28 allele, bilirubin levels, and coronary heart disease in the Framingham heart study. *Circulation* **114**, 1476–1481. https://doi.org/10.1161/CIRCULATIONAHA.106.633206 (2006).
66. Vítek, L. & Schwertner, H. A. The Heme catabolic pathway and its protective effects on oxidative stress-mediated diseases. *Adv. Clin. Chem.* **43**, 1–57. https://doi.org/10.1016/S0065-2423(06)43001-8 (2007).
67. Tang, L., Huang, C. & Feng, Y. Serum total bilirubin concentration is associated with carotid atherosclerosis in patients with prehypertension. *Clin. Exp. Hypertens.* **41**, 682–686. https://doi.org/10.1080/10641963.2018.1539094 (2019).
68. Vítek, L. The role of bilirubin in diabetes, metabolic syndrome, and cardiovascular diseases. *Front. Pharmacol.* **3**, 55. https://doi.org/10.3389/fphar.2012.00055 (2012).
69. Vitek, L. Bilirubin and atherosclerotic diseases. *Physiol. Res.* **66**, S11. https://doi.org/10.33549/physiolres.933581 (2017).
70. Kipp, Z. A. et al. Bilirubin levels are negatively correlated with adiposity in obese men and women, and its catabolized product, urobilin, is positively associated with insulin resistance. *Antioxidants* **12**, 170. https://doi.org/10.3390/antiox12010170 (2023).
71. Lee, M. J. et al. Serum bilirubin as a predictor of incident metabolic syndrome: a 4-year retrospective longitudinal study of 6205 initially healthy Korean men. *Diabetes Metab.* **40**, 305–309. https://doi.org/10.1016/j.diabet.2014.04.006 (2014).
72. Oda, E. & Aizawa, Y. Total bilirubin is inversely associated with metabolic syndrome but not a risk factor for metabolic syndrome in Japanese men and women. *Acta Diabetol.* **50**, 417–422. https://doi.org/10.1007/s00592-012-0447-5 (2013).
73. Li, X. H. et al. Direct bilirubin levels and risk of metabolic syndrome in healthy Chinese men. *Biomed. Res. Int.* **2017**, 9621615. https://doi.org/10.1155/2017/9621615 (2017).
74. Theodorakis, N. & Nikolaou, M. From Cardiovascular-Kidney-Metabolic syndrome to Cardiovascular-Renal-Hepatic-Metabolic syndrome: proposing an expanded framework. *Biomolecules* **15**, 213 (2025).
75. Shah, S. C. & Sass, D. A. Cardiac hepatopathy: a review of liver dysfunction in heart failure. *Liver Res. Open. J.* **1**, 1–10 (2015).
76. Liang, C. et al. Association of serum bilirubin with metabolic syndrome and non-alcoholic fatty liver disease: a systematic review and meta-analysis. *Front. Endocrinol. (Lausanne).* **13**, 869579 (2022).
77. Kim, A. H. et al. Sex differences in the relationship between serum total bilirubin and risk of incident metabolic syndrome in community-dwelling adults: propensity score analysis using longitudinal cohort data over 16 years. *Cardiovasc. Diabetol.* **23**, 92 (2024).
78. Song, Y., Yang, S. K., Kim, J. & Lee, D. C. Association between C-reactive protein and metabolic syndrome in Korean adults. *Korean J. Fam Med.* **40**, 116. https://doi.org/10.4082/kjfm.17.0075 (2019).
79. Oliveira, A. C. et al. RETRACTION: C-reactive protein and metabolic syndrome in youth: A strong relationship?? *Obesity* **16**, 1094–1098. https://doi.org/10.1038/oby.2008.43 (2008).
80. Yudkin, J. S. Cda. Stehouwer, jj. Emeis, sw. Coppack, C-reactive protein in healthy subjects: associations with obesity, insulin resistance, and endothelial dysfunction: a potential role for cytokines originating from adipose tissue? *Arterioscler. Thromb. Vasc Biol.* **19**, 972–978. https://doi.org/10.1161/01.ATV.19.4.972 (1999).
81. D'Alessandris, C., Lauro, R., Presta, I. & Sesti, G. C-reactive protein induces phosphorylation of insulin receptor substrate-1 on Ser 307 and Ser 612 in L6 myocytes, thereby impairing the insulin signalling pathway that promotes glucose transport. *Diabetologia* **50**, 840–849. https://doi.org/10.1007/s00125-006-0522-y (2007).
82. Akter, S., Shekhar, H. U. & Akhteruzzaman, S. Application of biochemical tests and machine learning techniques to diagnose and evaluate liver disease. *Adv. Bioscience Biotechnol.* **12**, 154–172 (2021).
83. Kwo, P. Y., Cohen, S. M. & Lim, J. K. ACG clinical guideline: evaluation of abnormal liver chemistries. *Official J. Am. Coll. Gastroenterology| ACG.* **112**, 18–35 (2017).

84. Chen, S. et al. Metabolic syndrome and serum liver enzymes in the general Chinese population. *Int. J. Environ. Res. Public. Health*. **13**, 223 (2016).
85. Zhang, H. et al. Machine learning-based prediction for 4-year risk of metabolic syndrome in adults: a retrospective cohort study. *Risk Manag. Healthc. Policy* **14**, 4361–4368. https://doi.org/10.2147/RMHP.S328180 (2021).

Acknowledgements
We extend our heartfelt appreciation to the data collection team and all participants whose contributions were crucial to the success of this study.

Author contributions
Bahareh Behkamal: Conceptualization, Data curation, Formal analysis.Fatemeh Asgharian Rezae: Writing – Original Draft.Amin Mansoori: Supervision, Writing – Review & Editing, Corresponding author.Rana Kolahi Ahari: Conceptualization, Data interpretation.Sobhan Mahmoudi Shamsabad: Methodology, Implementation. Mohammad Reza Esmaeilian: Methodology, Implementation.Gordon Ferns: Review & Editing.Mohammad Reza Saberi: Supervision, Corresponding author.Habibollah Esmaily: Supervision, Corresponding author.Majid Ghayour-Mobarhan: Review & Editing.

Declarations

Ethics approval and consent to participate
All study participants provided written informed consent prior to inclusion in the study. The research protocol and all procedures were reviewed and approved by the Ethics Committee of Mashhad University of Medical Sciences (Approval No. IR.MUMS.REC.1386.250), with ethical clearance granted in July 2007. All methods were performed in accordance with the relevant guidelines and regulations.

Consent for publication
Written informed consent for participation was obtained from all participants. Consent for image publication is not applicable to this manuscript, as all figures were developed only for the purpose of this study and do not involve any identifiable human images.

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to A.M., M.R.S. or H.E.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.