Development of A Robust Model for The Detection of Coronary Heart Disease Using Machine Learning Techniques

Abdulrahman Olalekan Yusuf

Ph.D. Student

Department of Electrical and Computer Engineering, Western Michigan University, USA

Graduate member, IEEE

May, 2025

Outline

- Introduction
- Problem statement
- Aim and objectives
- Significance of the work
- Methodology
- Results and Discussion
- Conclusion and Future work

Introduction

- Coronary heart disease (CHD) (also referred to as coronary artery disease or ischemic heart disease) is a heart disease that arises when the arteries of the heart cannot deliver enough oxygen-rich blood to the heart due to narrowing from the build-up of fatty deposits called plaque [1-3].
- Causes: excessive alcohol consumption, regular smoking, elevated cholesterol level, hypertension, and diabetes mellitus [4].
- CHD can be diagnosed using invasive angiography, computed tomography (CT) scan, or a 12-Lead Electrocardiogram (ECG) monitoring device [4].
- Machine learning (ML) and deep learning (DL) techniques are widely applied in biomedical care, healthcare, and disease prediction for non-invasive diagnosis of various cardiovascular diseases (CVD) [5-7].

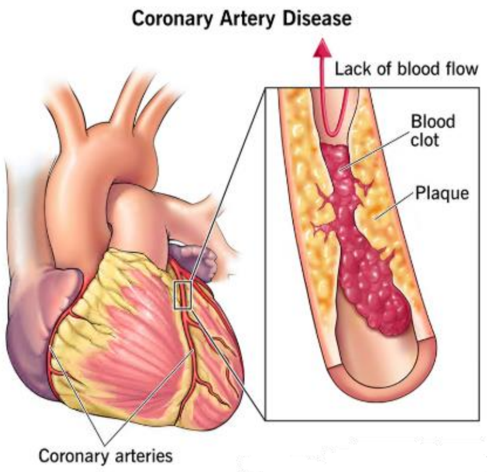


Fig. 1. Pictorial view of the coronary heart disease

Problem statement

- According to the Global Burden of Disease (GBD) study in 2013, the chronic disease group has a high mortality rate [4] and is frequently found in high-income countries such as the United States, which has 87% of deaths caused by chronic diseases [8].
- The conventional method of diagnosis of coronary heart disease (CHD) is very expensive, with different side effects, and requires strong technological knowledge.
- This study proposed the development of a robust model for the detection of coronary heart disease using machine learning (ML) techniques.

Emily Orta suffered a sudden cardiac arrest during a practice drill at the age of 14.

Source: [Sudden cardiac arrest in athletes: Prevention and management]

Aim and objectives of the work

The work aims to develop a robust model that can detect coronary heart disease using machine learning techniques.

The objectives are to:

- Process and clean the coronary heart disease public dataset.
- Compute and interpret the correlations of all the features.
- Apply validation set cross-validation technique to train and test different ML models to select the best ML model.
- Apply the k-fold cross-validation technique to tune the best model hyperparameters for optimal performance.
- Evaluate the performance of the model.

Significance of the work

- Low-cost and non-invasive approach for the diagnosis of coronary heart disease (CHD).
- Early detection of coronary heart disease for quick treatment.
- Aid medical personnel in the quick diagnosis of CHD.
- The developed model can be integrated into a wearable heart monitoring device [9] for a patient with a hereditary trait of CHD.

Methodology: CHD dataset

- Public Dataset: CVD_FinalData.csv [10]
- Number of observations: 5390
- Predictors: 21
  - Numerical data: 12 (Age, cigsPerDay, totChol, sysBP, diaBP, BMI, heartRate, glucose, Triglycerdie, hdl_cholesterol, ldl_cholesterol, CPK_MB_Percentage)
  - Categorical Data: 9
    - Binary category: 7 (sex, is_smoking, BPMeds, prevalentStroke, prevalentHyp, diabetes, exng)
    - Ordinal category: 2 (education, caa)
- Response or Target Variable: TenYearCHD (0 or 1) (Binary Classification)

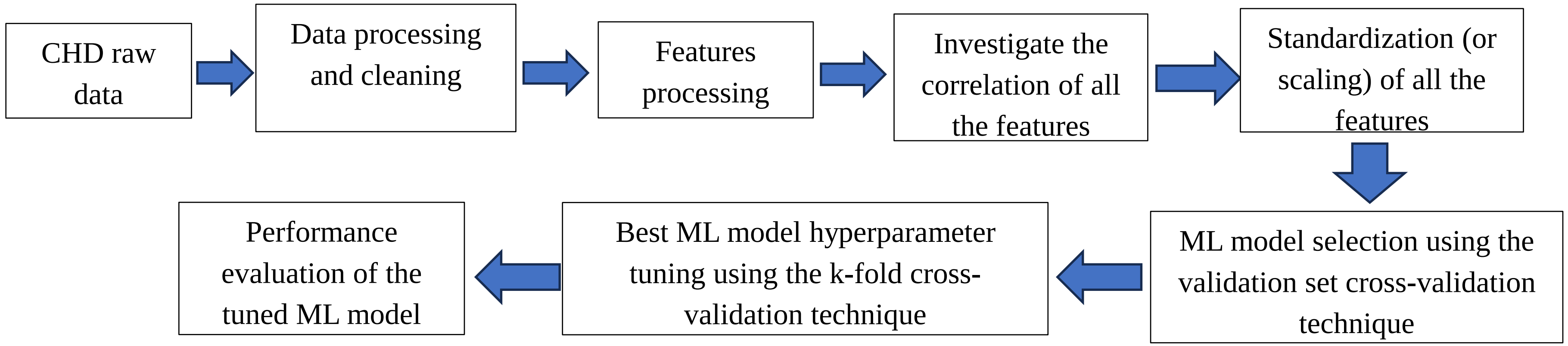
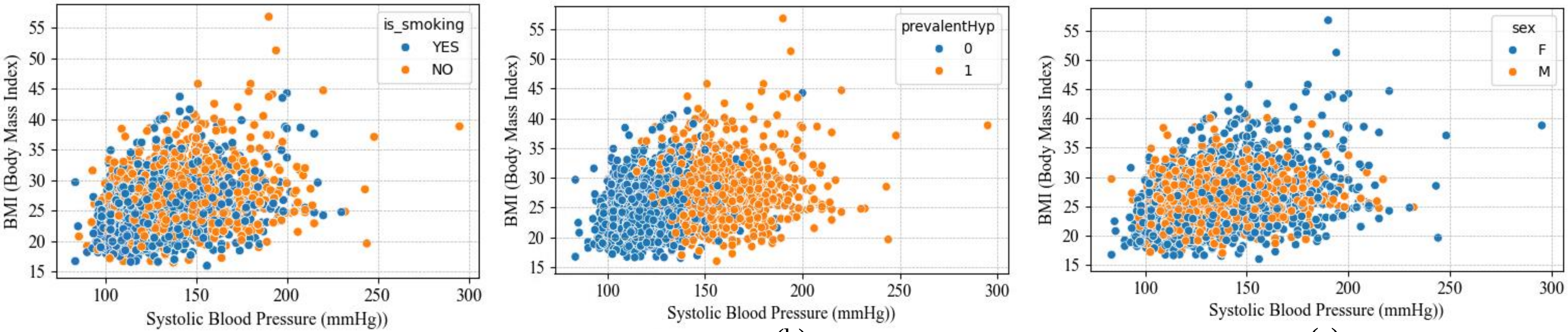


Fig. 2. Block diagram to develop a robust model for the detection of coronary heart disease using machine learning techniques

Methodology: CHD data visualization

Predictors: 21 (age (years), cigsPerDay (Cigarettes Per Day), totChol (Total Cholesterol (mg/dL)), sysBP (Systolic Blood Pressure (mmHg)), diaBP (Diastolic Blood Pressure (mmHg)), BMI (Body Mass Index), heartRate (bpm), glucose (Fasting Blood Glucose Level (mg/dL)), Triglycerdie (Level (mg/dL)), hdl_cholesterol (High-Density Lipoprotein Cholesterol), ldl_cholesterol (Low-Density Lipoprotein Cholesterol), CPK_MB_Percentage (Creatine Phosphokinase-MB Percentage (cardiac enzyme marker)), sex (M or F), is_smoking (Yes or No), BPMeds (Blood Pressure Medications (0 or 1)), prevalentStroke (History of Stroke (0 or 1)), prevalentHyp (History of Hypertension (High Blood Pressure)), diabetes (0 or 1), exng (Exercise-induced Angina (0 or 1)), caa (Number of Major Coronary Arteries Colored by Fluoroscopy (coronary artery anomaly/assessment)), education (degree level (1 - 4)).



(a) (b) (c)

Fig. 3. Relationship between the predictors: (a) BMI vs sysBP vs is_smoking (b) BMI vs sysBP vs prevalentHyp (c) BMI vs sysBP vs sex

Methodology: CHD data visualization cont'd

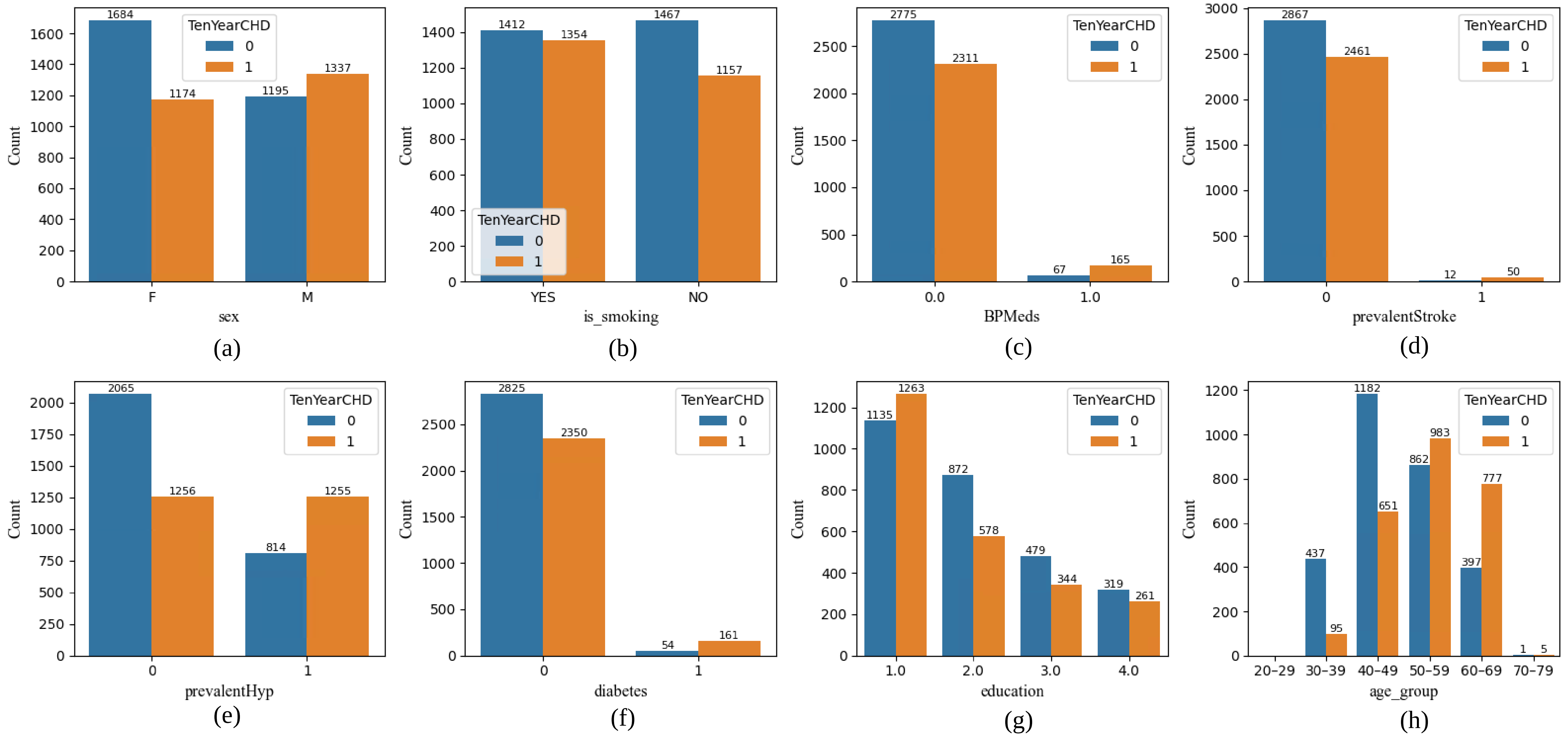
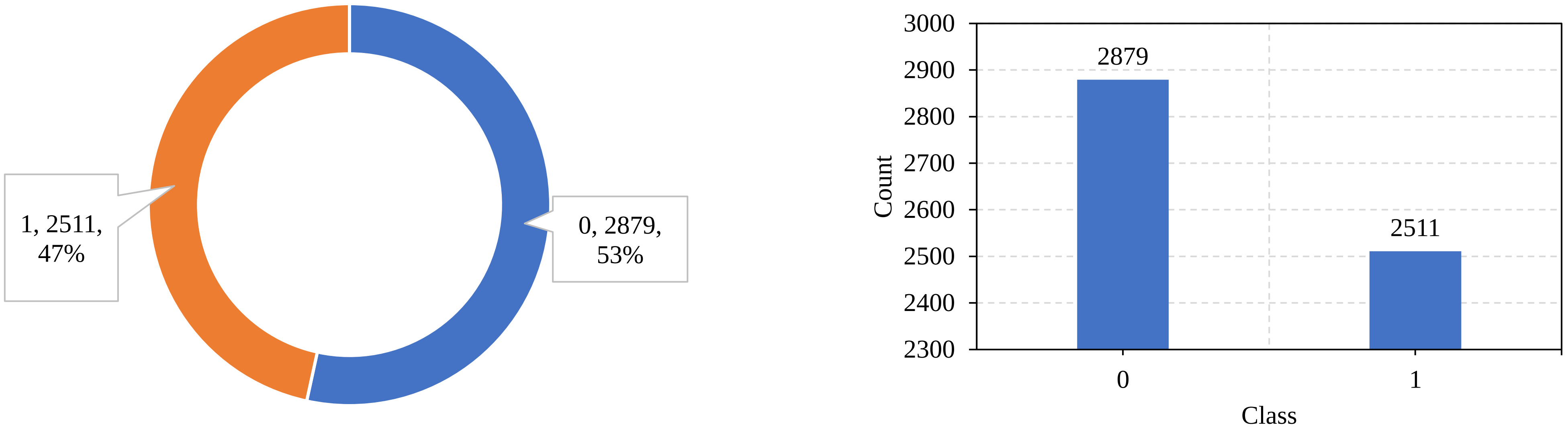


Fig. 4. Data distribution of the predictors: (a) sex (b) is_smoking (c) BPMeds (d) prevalentStroke (e) prevalentHyp (f) diabetes (g) education (h) age

Methodology: CHD data visualization cont’d

Response or Target Variable: TenYearCHD (0 or 1) (10-Year risk of CHD) (Binary indicator or risk probability)



(a) (b)

Fig. 5. Data summary of the target variable: (a) class distribution (b) data distribution

Methodology: Data cleaning and processing

- Fill in the missing values of the numerical data with the mean using simpleimputer pipeline.
- Fill in the missing values of the categorical data with the most frequent (mode) using simpleimputer pipeline.
- Convert all the categorical data to numerical values using the binary encorder technique

	age	education	cigsPerDay	BPMeds	prevalentStroke	prevalentHyp	diabetes	totChol	sysBP	diaBP	...	glucose	exng	caa	Triglyceride	hdl_cholestrol	ldl_cholestrol	CPK_MB_Percentage	sex	is_smoking	TenYearCHD
0	64.0	2.0	3.0	0.0	0.0	0.0	0.0	221.0	148.0	85.0	...	80.0	1.0	1.0	212.0	32.0	173.0	5.0	F	YES	1
1	36.0	4.0	0.0	0.0	0.0	1.0	0.0	212.0	168.0	98.0	...	75.0	0.0	0.0	112.0	35.0	124.0	4.0	M	NO	0
2	46.0	1.0	10.0	0.0	0.0	0.0	0.0	250.0	116.0	71.0	...	94.0	0.0	0.0	135.0	60.0	126.0	5.0	F	YES	0
3	50.0	1.0	20.0	0.0	0.0	1.0	0.0	233.0	158.0	88.0	...	94.0	1.0	2.0	232.0	28.0	200.0	7.0	M	YES	1
4	64.0	1.0	30.0	0.0	0.0	0.0	0.0	241.0	136.5	85.0	...	77.0	0.0	0.0	143.0	38.0	111.0	5.0	F	YES	0

[5 rows x 22 columns]
PS C:\Users\CEAS User\Desktop\IEEE Dataport>

(a)

	age	education	cigsPerDay	BPMeds	prevalentStroke	prevalentHyp	diabetes	totChol	sysBP	diaBP	...	glucose	exng	caa	Triglyceride	hdl_cholestrol	ldl_cholestrol	CPK_MB_Percentage	sex	is_smoking	TenYearCHD
0	64.0	2.0	3.0	0.0	0.0	0.0	0.0	221.0	148.0	85.0	...	80.0	1.0	1.0	212.0	32.0	173.0	5.0	0	1	1
1	36.0	4.0	0.0	0.0	0.0	1.0	0.0	212.0	168.0	98.0	...	75.0	0.0	0.0	112.0	35.0	124.0	4.0	1	0	0
2	46.0	1.0	10.0	0.0	0.0	0.0	0.0	250.0	116.0	71.0	...	94.0	0.0	0.0	135.0	60.0	126.0	5.0	0	1	0
3	50.0	1.0	20.0	0.0	0.0	1.0	0.0	233.0	158.0	88.0	...	94.0	1.0	2.0	232.0	28.0	200.0	7.0	1	1	1
4	64.0	1.0	30.0	0.0	0.0	0.0	0.0	241.0	136.5	85.0	...	77.0	0.0	0.0	143.0	38.0	111.0	5.0	0	1	0

[5 rows x 22 columns]
PS C:\Users\CEAS User\Desktop\IEEE Dataport>

(b)

Fig. 6. CHD data cleaning and processing: (a) filling the missing values (b) binary encoding of the categorical data

Methodology: Feature correlation

- Observe the relationship between the features [11]
- Range: -1 to +1
  - +1: strong positive relationship (i.e increase in a variable tends to increase the other variable)
  - 0: No relationship
  - -1: strong negative relationship (i.e increase in a variable tends to decrease the other variable)
- Numerical vs Numerical: Pearson correlation technique
- Numerical vs Binary Categorical: point-biserial correlation technique
- Numerical vs Ordinal Categorical: Spearman correlation technique
- Binary Categorical vs Binary Categorical: Phi coefficient (ϕ) correlation technique
- Binary Categorical vs Ordinal Categorical: Spearman correlation technique
- Ordinal Categorical vs Ordinal Categorical: Spearman correlation technique

Methodology: Machine learning classifiers

- ML classifiers [12]:
  - K-nearest neigbors (KNN)
  - Support vector machine (SVM)
  - Decision tree (DT)
  - Random forest (RF)
  - Logistic regression (LR)
  - Gaussian naïve bayes (GNB)
  - Linear discriminant analysis (LDA)
  - Light gradient boost (LGBoost)
  - Extreme gradient boost (XGBoost)
  - CatBoost (CBoost)

- Standardization (scaling) of all the features to ensure all the features contribute equally to the model and help to reduce bias.
- Model selection: validation set cross-validation technique: (80%-20 %), 80% dataset for training and 20% dataset for testing the ML model.
- K-fold cross-validation technique: 10-folds of the training dataset to tune the hyperparameter of the best model

Methodology: Validation set cross-validation technique (80%-20%)

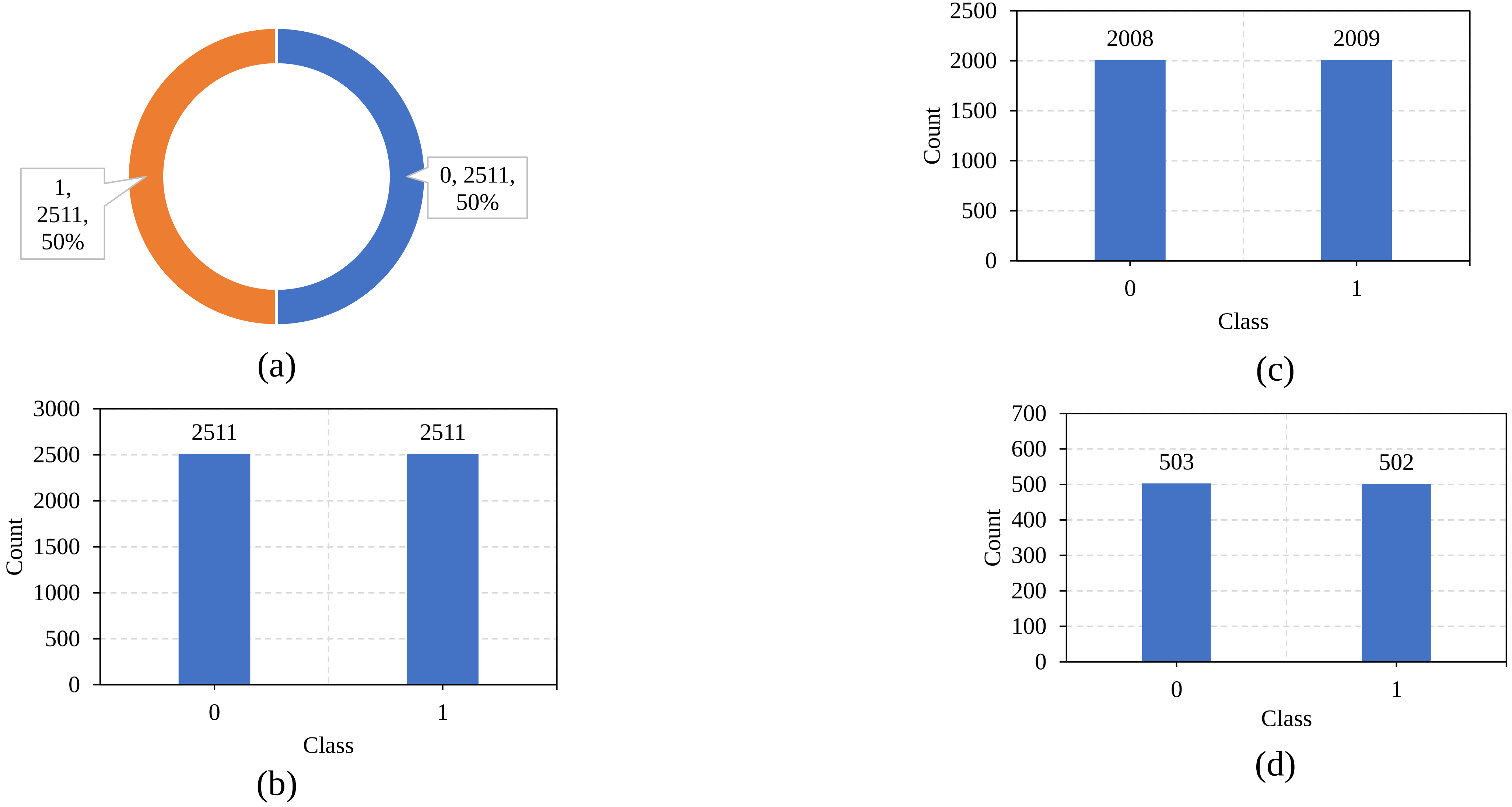


Fig. 7. Validation set cross-validation (80%-20%): (a) balanced class dataset (b) data distribution (c) training dataset (d) test dataset

Methodology: ML model performance metrics

Performance metrics [12]:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Sensitivity\ (Recall) = \frac{TP}{TP+FN} \quad (2)$$

$$Specificity = \frac{TN}{TN+FP} \quad (3)$$

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

$$F1 - score = \frac{2\times Sensitivity \times Precision}{Sensitivity + Precision} \quad (5)$$

$$Geometric\ mean\ (G - mean) = \sqrt{Sensitivity \times Specificity} \quad (6)$$

$where, TP - True\ Positive, TN - True\ Negative, FP - False\ Positive, FN - False\ Negative$

Results and discussion: Correlation of all the features

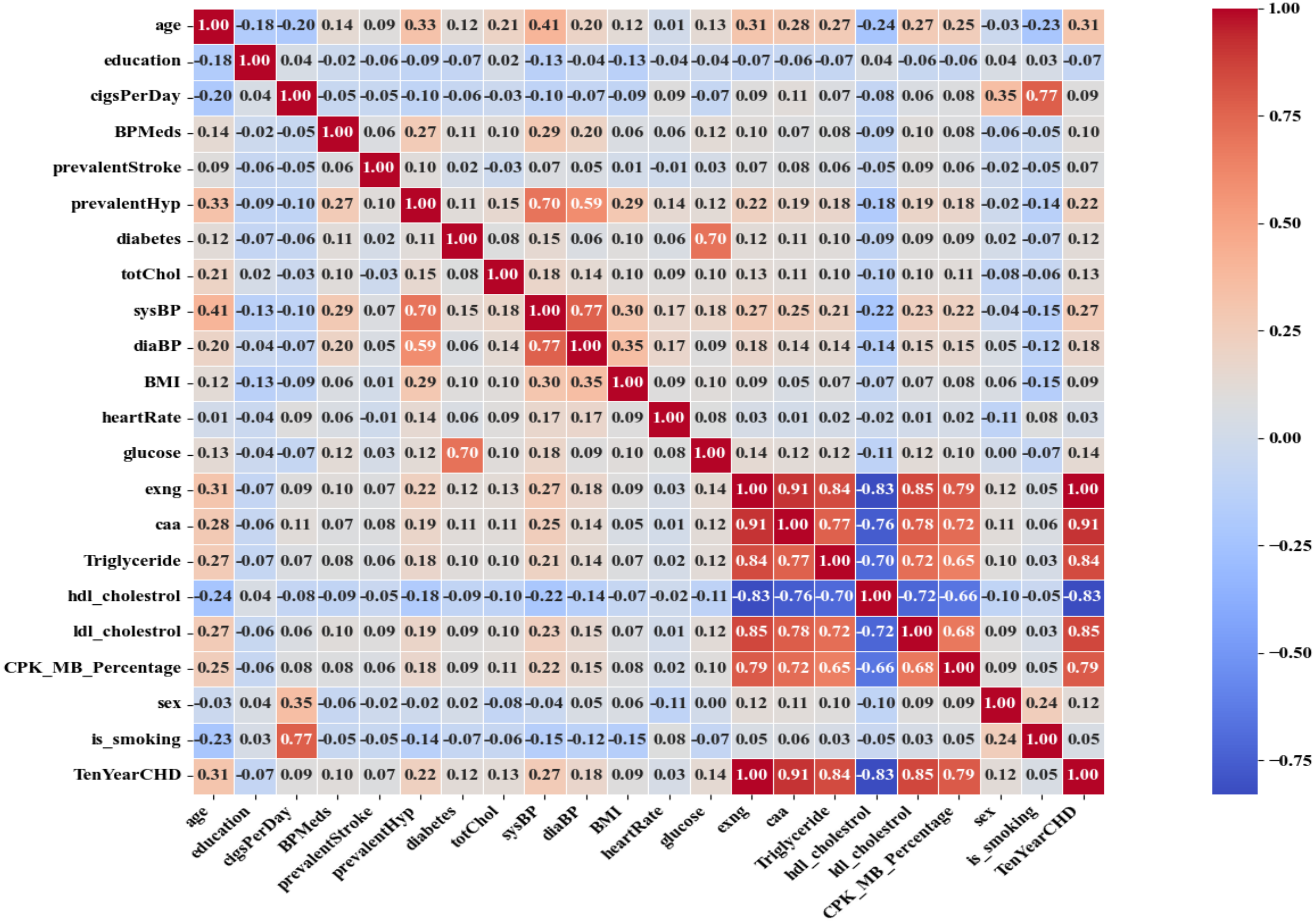


Fig. 8. Correlation matrix of all the features

Results and discussion: Correlation of the top features cont'd

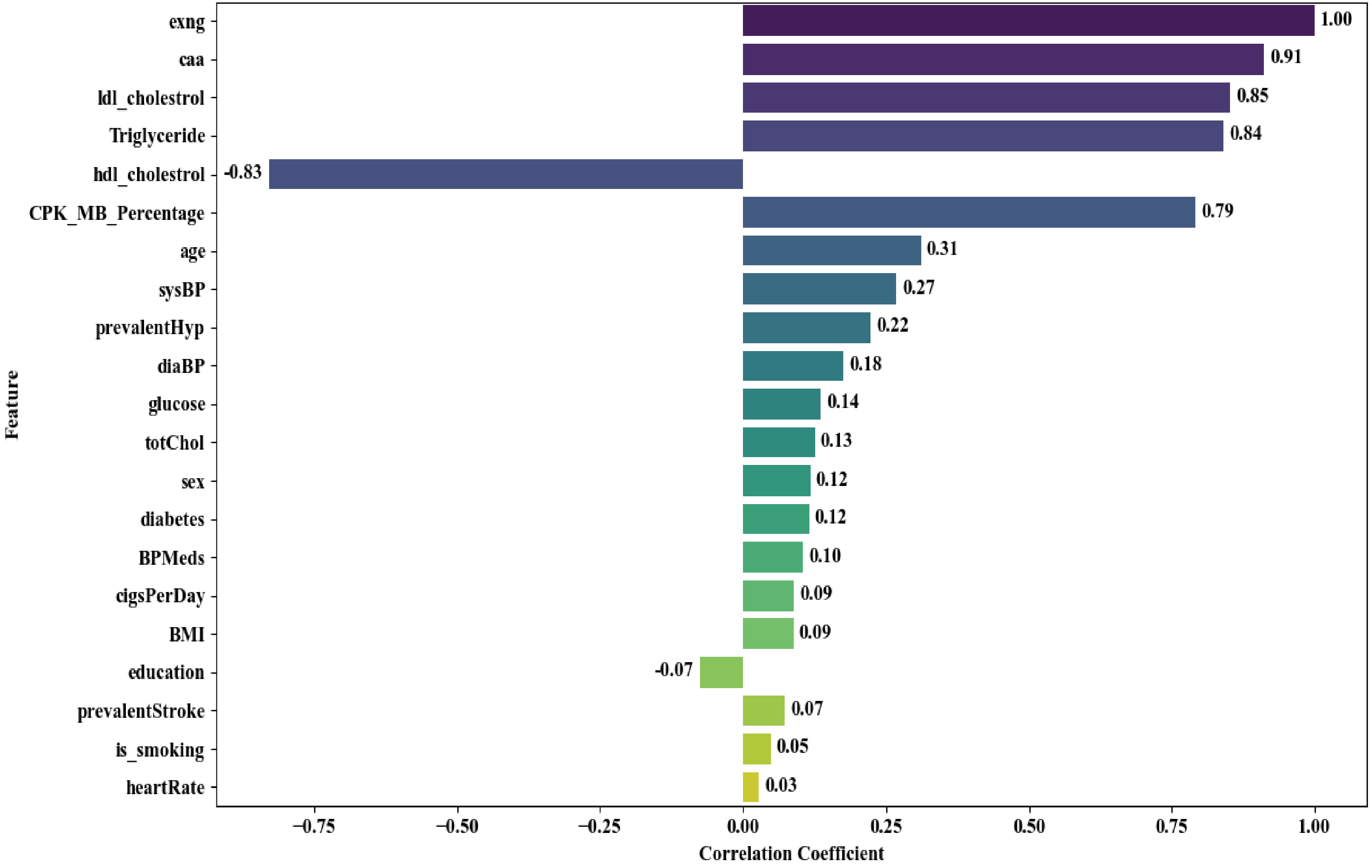


Fig. 9. Top correlated features with the target variable (TenYearCHD)

Comments:

- The exng, caa, ldl_cholesterol, Triglyceride, hdl_cholestrol and CPK_MP_Percentage features show strong correlation to coronary heart disease (CHD)
- The exng, caa, ldl_cholesterol, Triglyceride, and CPK_MP_Percentage features have a strong positive correlation (> 0.5) to CHD (i.e the higher the value of the features, the higher CHD)
- The hdl_cholestrol feature has a strong negative correlation (< -0.5) to CHD (i.e, the higher the value of the feature, the lower the CHD)

Results and discussion: ML model selection with different classifiers

Table 1. ML model selection using different ML classifiers with 21 predictors

S/No.	Machine learning classifier	Accuracy	Sensitivity	Specificity	Precision	F1-score	G-mean
1	Random Forest	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	Linear Discriminant Analysis	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3	K-nearest neigbor	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	Support vector machine	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5	Decision tree	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6	Gaussian naïve bayes	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
7	Logistic regression	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
8	Light gradient boost	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
9	Extreme gradient boost	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
10	Catboost	1.0000	1.0000	1.00000	1.0000	1.0000	1.0000

Comment: The exng, caa, ldl_cholesterol, Triglyceride, hdl_cholesterol, and CPK_MP_Percentage features show strong correlation to coronary heart disease (CHD), these predictors are binary indicators, which may cause the model to overfit, and model overfitting is suspected. However, if all the 21 predictors are considered, then all the ML classifier models would perform excellently.

***Drop predictors with strong correlation: exng, caa, ldl_cholesterol, Triglyceride, hdl_cholesterol, and CPK_MP_Percentage, then train and test ML model with the remaining 15 predictors.

Results and discussion: ML model selection with different classifiers cont’d

Table 2. ML model selection using different ML classifiers with 15 predictors

S/No.	Machine learning classifier	Accuracy	Sensitivity	Specificity	Precision	F1-score	G-mean
1	Random Forest	0.9741	0.9742	0.9742	0.9754	0.9748	0.9742
2	Linear Discriminant Analysis	0.6557	0.6557	0.6557	0.6569	0.6563	0.6557
3	K-nearest neigbor	0.7642	0.7643	0.7643	0.7762	0.7702	0.7643
4	Support vector machine	0.7065	0.7064	0.7064	0.7077	0.7071	0.7064
5	Decision tree	0.9045	0.9046	0.9046	0.9197	0.9121	0.9046
6	Gaussian naïve bayes	0.5841	0.5837	0.5837	0.6564	0.6179	0.5837
7	Logistic regression	0.6637	0.6636	0.6636	0.6650	0.6643	0.6636
8	Light gradient boost	0.9114	0.9115	0.9115	0.9184	0.9149	0.9115
9	Extreme gradient boost	0.9363	0.9364	0.9364	0.9430	0.9397	0.9364
10	Catboost	0.8667	0.8667	0.8667	0.8675	0.8671	0.8667

Comments: Selecting predictors with less correlation: (Age, cigsPerDay, totChol, sysBP, diaBP, BMI, heartRate, glucose, sex, is_smoking, BPMeds, prevalentStroke, prevalentHyp, diabetes, education)

Random forest has the best performance with test accuracy, test sensitivity, test specificity, test precision, test F1-score, and test G-mean of 97.41%, 97.42%, 97.42%, 97.54%, 97.48%, and 97.42%, respectively.

Results and discussion: ML model hyperparameter tuning

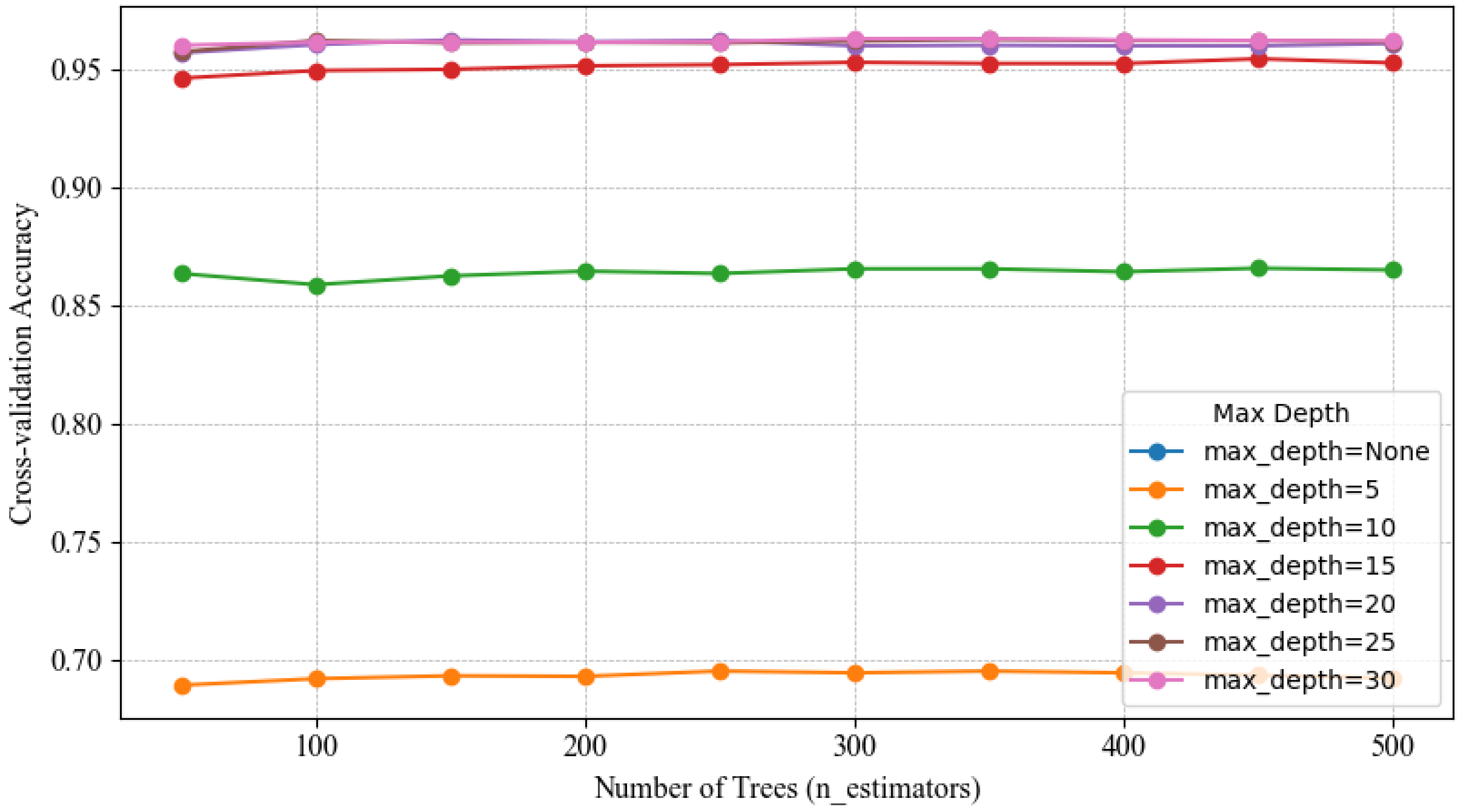


Fig. 10. Hyperparameter tuning using the k-fold cross-validation technique (10-fold) to optimize the random forest classifier performance (Number of trees: 300, max_depth: None, and cross-validation accuracy: 96.29%)

Performance metrics:

- Test accuracy: 97.71%
- Test sensitivity: 97.71%
- Test specificity: 97.81%
- Test precision: 97.81%
- Test F1-score: 97.76%
- Test G-mean: 97.71%

Results and discussion: ML model performance evaluation

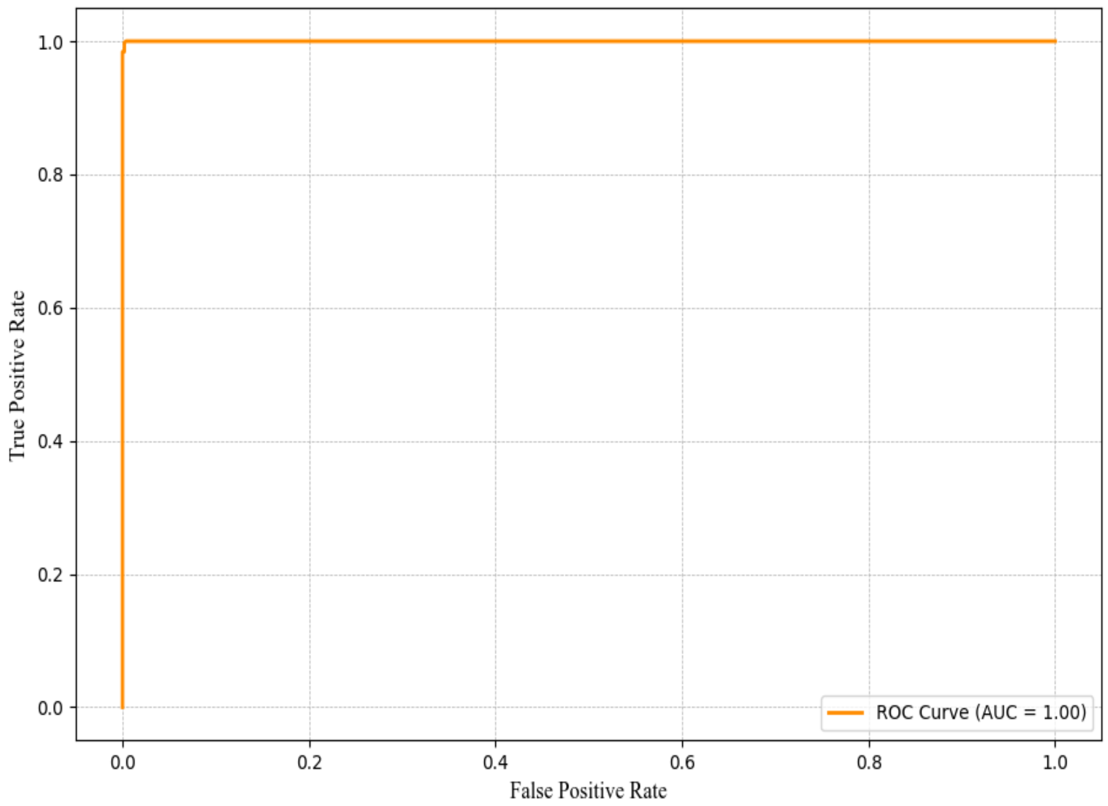


Fig. 11. Area under the Receiver Operating Characteristic (AUROC) curve for the random forest model to evaluate potential overfitting and assess classifier performance across all thresholds

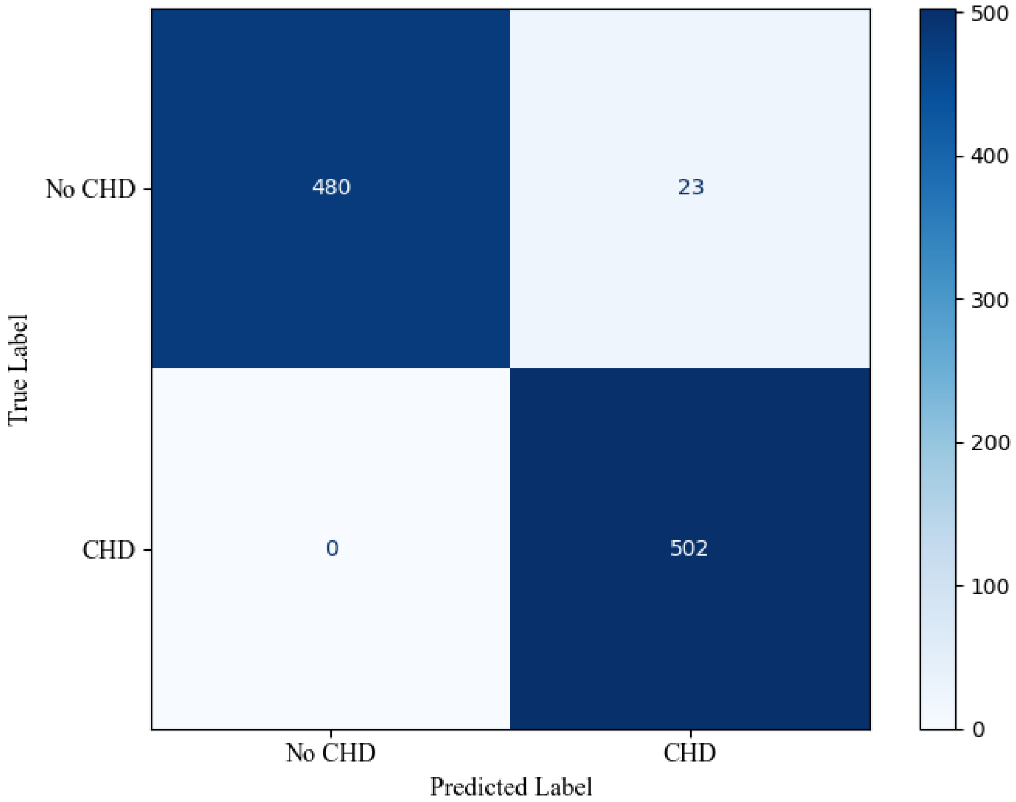


Fig. 12. Confusion matrix of the classes using the random forest ML model

The AUROC of the random forest model is 1.00, which signifies the trained model can clearly distinguish between the two classes at every threshold, which is very good and signifies the model can distinctly classify the two classes.

Results and discussion: Results comparison with the literature

Table 3. Comparing the results with the literature

S/No.	Authors	Methodology	Findings	Limitation	Reference
1	R. Yılmaz and F. H. Yağın, 2022	• Early detection of CHD using ML methods • Dataset: IEEEdataPort from Cleveland, Hungarian, Switzerland, Statlog (Heart) Data Set, and Long Beach VA (F: 281, M: 909, Total: 1,190 observations) • Number of predictors: 11 • ML model: LR, SVM, RF	• Classifier accuracy: RF: 92.9%, 89.7%, 86.1%, RF has the best performance • Accuracy, specificity, sensitivity, F1-score, negative predictive and positive predictive values of RF are 92.9%, 92.8%, 92.8%, 92.9%, and 92.8% respectively.	• Considered only 3 ML models • Accuracy can be improved • No test for ML model overfitting	[4]
2	V. Shorewala, 2021	• Early detection of CHD using ensemble ML techniques • Dataset: 2019 Cardiovascular Disease Dataset (70,000 observations) from Kaggle • Number of predictors: 11 • ML model: Bagging (RF, DT), boosting (LGBoost, XGBoost), and stacking ML classifiers and Neural network	• Neural network accuracy: 73.82% • Bagging and boosting ML classifiers have an accuracy of 73.82% and 73.89%, respectively. • Stacking ensemble technique (KNN, RF, SVM) is effective with an accuracy of 75.1% .	• Challenging to tune the hyperparameters of the stacking ensemble technique. • Accuracy can be improved.	[13]
3	Hassan *et. al.*, 2022	• Effectively predicting CHD using ML classifiers • Dataset: University of California Irvine (UCI-repository) (303 observations) • Number of predictors: 13 • Validation set (70% for training and 30% for testing) • ML model: kNN, RF, gradient boosted tree (GBT), Multilayer Perceptron (MLP) (neural network)	• kNN has the worst performance, while GBT and MLP have accuracy of 95%. • RF model has the best performance with an accuracy of 96% in heart disease prediction.	• Considered only 4 ML models • kNN performance accuracy was not stated. • Accuracy can be improved • No test for ML model overfitting • Small dataset	[14]
4	Proposed model	• Classification of CHD using ensemble ML models • Dataset: 2025 IEEEdataPort (F: 2,858, M: 2,532, Total: 5,390 observations) • Number of predictors: 15 • ML model: LR, SVM, RF, DT, KNN, GNB, LDA, LGBoost, XGBoost, CBoost	• RF model has the best performance with accuracy, sensitivity, specificity, precision, F1-score, and G-mean of 97.41%, 97.42%, 97.42%, 97.54%, 97.48%, and 97.42%, respectively. • The hyperparameters of the RF model were tuned to number of trees of 300 and none max_depth to obtain test accuracy, sensitivity, specificity, precision, F1-score, G-mean and AUROC of 97.71%, 97.71%, 97.81%, 97.54%, 97.76%, 97.71%, and 1.00 respectively.	• Accuracy can be improved • Consider more observations by combining similar datasets. • Explore stacking ensemble ML technique and deep learning techniques (Neural network)	

Conclusion

- The public dataset for ten years of coronary heart disease (CHD) (CVD_FinalData) was processed to train and test a robust ML model for predicting the risk of coronary heart disease.
- The exng, caa, ldl_cholesterol, Triglyceride, hdl_cholestrol and CPK_MP_Percentage features show strong correlation to the CHD. The exng, caa, ldl_cholesterol, Triglyceride, and CPK_MP_Percentage features have a strong positive correlation to CHD, while the hdl_cholestrol feature has a strong negative correlation to CHD.
- Training and testing of the ML classifier with all the 21 features yielded test accuracy, sensitivity, specificity, precision, F1-score, and G-mean of 100.00%, 100.00%, 100.00%, 100.00%, 100.00%, and 100.00% respectively which signifies all the classifiers can perfectly classify the target variable but model overfitting is suspected and the predictors with strong correlation were dropped that may cause model overfitting.
- Different ML classifiers were trained and tested considering 15 features of the dataset using the validation set cross-validation technique to prevent overfitting and select the best model. The random forest model performs best with test accuracy, sensitivity, specificity, precision, F1-score, and G-mean of 97.41%, 97.42%, 97.42%, 97.54%, 97.48%, and 97.42%, respectively.
- The hyperparameters of the random forest were tuned with a number of trees of 300 and none max_depth to obtain test accuracy, sensitivity, specificity, precision, F1-score, and G-mean of 97.71%, 97.71%, 97.81%, 97.54%, 97.76%, and 97.71%, respectively. Furthermore, the AUROC of the random forest model is 1.00 for all the classes, which signifies the ML model can distinctly classify the two classes across all thresholds using the 15 predictors.
- A random forest model was trained and tuned with optimum performance using machine learning techniques to predict the risk of coronary heart disease.
- The code and dataset used in the work are available at https://github.com/abdulrahman28/IEEE_Dataport_AI_Solution
- The video for the presentation is available at https://drive.google.com/drive/folders/1wyJuLy0LQ-w8rEGb25UJdK45PjAVrNfa?usp=drive_link

Future work

- Explore more publicly available CHD datasets that can be combined to have a large dataset for implementing deep learning techniques such as neural networks, to improve model performance.
- ML model integration into wearable heart monitoring devices to notify the patient or medical personnel if CHD is suspected.

REFERENCES

[1] Y. Miao, Y. Wang, and Q. Wan, "Association between TyG index with obesity indicators and coronary heart disease: a cohort study," *Scientific Reports*, vol. 15, p. 8920, 2025.

[2] C.-Y. Ma, Y.-M. Luo, T.-Y. Zhang, Y.-D. Hao, X.-Q. Xie, X.-W. Liu, *et al.*, "Predicting coronary heart disease in Chinese diabetics using machine learning," *Computers in Biology and Medicine*, vol. 169, p. 107952, 2024.

[3] B. Ramesh and K. Lakshmanna, "A novel early detection and prevention of coronary heart disease framework using hybrid deep learning model and neural fuzzy inference system," *IEEE Access*, vol. 12, pp. 26683-26695, 2024.

[4] R. Yılmaz and F. H. Yağın, "Early detection of coronary heart disease based on machine learning methods," *Medical Records*, vol. 4, pp. 1-6, 2022.

[5] T. Ullah, S. I. Ullah, K. Ullah, M. Ishaq, A. Khan, Y. Y. Ghadi, et al., "Machine learning-based cardiovascular disease detection using optimal feature selection," *IEEE Access*, vol. 12, pp. 16431-16446, 2024.

[6] K. Baseer, K. Sivakumar, D. Veeraiah, G. Chhabra, P. K. Lakineni, M. J. Pasha, et al., "Healthcare diagnostics with an adaptive deep learning model integrated with the Internet of medical Things (IoMT) for predicting heart disease," *Biomedical Signal Processing and Control*, vol. 92, p. 105988, 2024.

[7] A. Ogunpola, F. Saeed, S. Basurra, A. M. Albarrak, and S. N. Qasem, "Machine learning-based predictive models for detection of cardiovascular diseases," *Diagnostics*, vol. 14, p. 144, 2024.

[8] S. B. Patil and Y. Kumaraswamy, "Intelligent and effective heart attack prediction system using data mining and artificial neural network," *European Journal of Scientific Research*, vol. 31, pp. 642-656, 2009.

[9] M., Simin, M. Panahi, D. Maddipatla, A. J. Hanson, S. Fenech, L. Bonek, N. Sapoznik, P. D. Fleming, B. J. Bazuin, and M. Z. Atashbar. "Development of a flexible wireless ECG monitoring device with dry fabric electrodes for wearable applications," IEEE Sensors Journal, vol. 22, pp. 11223-11232, 2021.

[10] Ye Chen, Bang Li, Yaoting Zhang, Yang Zheng, He Cai, Muhammad Faisal, "Wearable Health Hackathon - Develop a solution to early diagnose diseases using non-invasive methods and health data", *IEEE Dataport*, March 16, 2025, doi:10.21227/vne2-zz58.

[11] K. Okoye and S. Hosseini, "Correlation tests in R: pearson cor, kendall’s tau, and spearman’s rho," in R programming: Statistical data analysis in research, ed: Springer, 2024, pp. 247-277.

[12] G. Naidu, T. Zuva, and E. M. Sibanda, "A review of evaluation metrics in machine learning algorithms," in Computer science on-line conference, 2023, pp. 15-25.

[13] V. Shorewala, "Early detection of coronary heart disease using ensemble techniques," Informatics in Medicine Unlocked, vol. 26, p. 100655, 2021.

[14] C. A. U. Hassan, J. Iqbal, R. Irfan, S. Hussain, A. D. Algarni, S. S. H. Bukhari, N. Alturki, and S. S. Ullah, "Effectively predicting the presence of coronary heart disease using machine learning classifiers," Sensors, vol. 22, p. 7227, 2022.

Thank you!!!

Abdulrahman Olalekan Yusuf

abdulrahmanolalekan.yusuf@wmich.edu

Acknowledgments:
This research was supported by the Department of Electrical and Computer Engineering and funded by a Faculty Research and Creative Activities Award from Western Michigan University.