



Explainable Machine Learning on CAD-RADS Score Classification Based on Heart Disease Risk Factors

Jondri^{*}, Indwiarti^{}, Diyas Puspandari^{}

School of Computing, Telkom University, Bandung 40257, Indonesia

Corresponding Author Email: jondri@telkomuniversity.ac.id

Copyright: ©2025 The authors. This article is published by IIETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/isi.300316>

ABSTRACT

Received: 16 November 2024

Revised: 25 January 2025

Accepted: 27 March 2025

Available online: 31 March 2025

Keywords:

CAD-RADS score, multi-layer perceptron, explainable machine learning, SHAP, LIME, box plot

Machine learning performance in classification has significantly improved, fueled by the development of more advanced learning algorithms and enhanced by faster computational speeds. In several cases, artificial intelligence has beaten a panel of doctors in classifying data. However, there is a weakness in this machine learning algorithm, namely that it cannot explain why data falls into a certain class. For cases in the world of health, this explanation is much more important than simply judging that someone has a certain disease. This study uses a multi layer perceptron (MLP) as the classification method to explore the connection between risk factors for heart disease and coronary artery disease reporting and data system (CAD-RADS) scores. Analysis of influential features (global analysis) using SHapley Additive explanations (SHAP), while local explanations using local interpretable model-agnostic explanations (LIME). This paper also conducted an in-depth exploratory data analysis (EDA) on the features using box plots. The best classification result with an accuracy value 0.87 was obtained using the symptom and examination feature groups. Meanwhile, if viewed from the F1 score and AUC values, the best features are all features with an F1 score of 0.83 and an AUC value of 0.92.

1. INTRODUCTION

Coronary artery disease (CAD) is when the coronary artery series, which supply blood to the heart, become narrowed or blocked due to fatty deposits, inflammation, or other factors. This blockage can lead to a heart attack or stroke [1].

Diagnosis of CAD can be done invasively or non-invasively. Invasive diagnostic methods involve procedures that directly visualize the coronary arteries. The main invasive method for diagnosing CAD is Coronary Angiography. This involves injecting a contrast agent into the coronary arteries through a catheter inserted through an artery in the leg or arm. X-ray images are then taken to visualize the coronary arteries and detect any blockages or narrowing of the arteries [2].

Non-invasive methods involve testing that does not require the insertion of instruments into the body. One common non-invasive diagnostic technique for CAD uses a computed tomography (CT) scan to measure calcium deposits in the coronary arteries. Higher calcium scores indicate more significant plaque buildup and potential risk of CAD [3, 4]. Coronary CT angiography (CCTA) is a specialized imaging technique that uses a computed tomography (CT) scan to visualize the coronary arteries and assess for blockages or disease. CCTA typically uses a high-speed CT scanner with multiple detector rows to quickly and accurately capture images of the heart and coronary arteries [3].

Coronary angiography remains the gold standard for definitive diagnosis of CAD, while non-invasive tests offer valuable tools for early screening and risk assessment [5].

The CAD-RADS was developed to ensure consistency in reporting for patients receiving coronary CT angiography (CCTA) and to assist in determining appropriate follow-up actions for patient care [1]. CAD-RADS categorizes CAD patients based on the severity of stenosis and plaque thickness [6, 7]. CAD-RADS categories range from CAD-RADS 0, indicating no plaque, to CAD-RADS 5, representing at least one complete occlusion. Each category is defined based on the highest degree of stenosis present. For example, a single maximum stenosis of 1%–24% is classified as CAD-RADS 1, 25%–49% as CAD-RADS 2, 50%–69% as CAD-RADS 3, and 70%–99% as CAD-RADS 4. CAD-RADS 4 is further divided into subcategories 4A and 4B, depending on whether the stenosis is in one or two vessels or three vessels or the left main artery. CAD-RADS 5 represents complete occlusion of at least one vessel, which can be acute or chronic [7].

Using machine learning techniques to classify CAD-RADS based on Computed Tomography Coronary Angiography (CCTA) images has become an active area of research in recent years. This field combines advanced imaging techniques with sophisticated machine learning algorithms to improve diagnostic accuracy. Studies that have been conducted include using Convolutional Neural Network (CNN) for CAD-RADS score classification based on CCTA images [8–11]. In addition, there are also studies on heart disease risk factors that affect CAD-RADS scores. Machine learning methods used include Random Forest (RF), K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Neural Network (NN), Decision Tree Classification (DTC)

and Linear Discriminant Analysis (LDA) [12]. The machine learning model used is a black box model, both in classifying images into CAD-RADS scores and in determining heart disease risk factors that affect CAD-RADS scores. The relationship between input and output cannot be explained.

The novelty of this study is discussing explainable machine learning (EML) models that can explain the results of the black box models used. In addition, EDA was carried out using box plots to provide an overview of the distribution of data in the normal and CAD patient data classes.

2. METHODOLOGY

This research was conducted in 4 stages, namely data input, preprocessing, classification, and analysis, as seen in Figure 1.

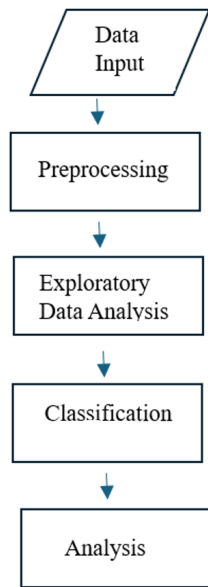


Figure 1. Research flow chart

The data is derived from secondary sources from the UCI Machine Learning Repository. The dataset used is the Z-Alizadeh Sani dataset, which comes from visitors to Shaheed Rajaei Cardiovascular, Medical and Research Center [13]. The data comprises 303 rows (216 CAD patients and 87 normal) with 55 features. These features are grouped into four categories, namely: (1) Demographic, (2) symptoms and examination, (3) electrocardiogram, and (4) laboratory and echo features. The data has two classes, namely class 0, which states normal people (not suffering from CAD), and class 1, which states data on CAD sufferers.

The preprocessing performed is the normalization of features that have a numerical value greater than 1, such as age, height, BMI, and others. So that all features have a value between 0 and 1.

To get an overview of the data distribution in class 0 and class 1 data, Exploratory Data Analysis (EDA) was carried out. The EDA tool used is a boxplot on features with continuous values. In order for the boxplots of features that have different values can be plotted on the same image, the boxplot is applied to the normalized data. This section discusses the measure of data centralization (median), the measure of position (quartile), and the measure of data spread (distance between quartiles). From the boxplot, it can also be seen whether there is data that differs from the others or

outliers.

The classification method used is MLP. The architecture of the multilayer perceptron used is two hidden layers with 5 neurons in the first hidden layer and 6 neurons in the second hidden layer. The activation function used is Relu. The parameters set are `random_state = 5`, `verbose = True`, `learning_rate_init = 0.01`, `max_iter = 4000`, `tol = 0.00000001`. While other parameters are defaulted from sklearn. As much as 80% of the data is used as training data, and the rest as testing data.

Several EML techniques will be used in the analysis stage. The results of the CAD-RADS score classification using MLP already have good performance, as measured by accuracy, F1 score, sensitivity, specificity or AUC. However, MLP cannot explain the influence of each feature on the data class [14]. In this study, two EML methods will be used: SHAP and LIME.

2.1 SHAP

SHAP aims to explain predictions from data by calculating the contribution of each feature to the output class. In the SHAP analysis process, the prediction for each instance is explained by assessing the contribution of each feature using SHapley values, which are derived through the application of game theory coalition methods [15, 16]. The SHapley value for each feature is calculated by averaging over all possible combinations of feature values. These values directly reflect the influence of each feature on the prediction. To quantify the overall impact of each feature globally, the SHapley values for all features in the dataset are averaged. Finally, the feature values are sorted in descending order of importance, and a plot is generated to visualize the results [17, 18].

2.2 LIME

As the name implies, LIME is an EML method that can provide an explanation of the influence of features on certain data [19]. This is done by using a model that can be interpreted against a data. LIME begins by selecting one particular data to be explained. Next, a new dataset is created by randomly generating data that is around the data. The class of data that has been generated is obtained from the prediction of the black box model that has been trained on the training data. This new dataset is weighted based on its proximity to the data to be interpreted. On this new dataset, a new interpretable model is trained [20].

3. RESULTS AND DISCUSSION

3.1 EDA

In this section, a descriptive analysis of the data using boxpot will be conducted. The analysis is conducted on the demographic, laboratory, and symptoms and examination data categories. The features that are made into box plots are features that have continuous numeric values. Because the data has been normalized, the box plot of the data in the same category can be plotted on the same image. So that the data distribution of a feature in the normal class (0) and the CAD class (1) can be compared.

In Figure 2 and Table 1, it can be seen that the age feature has a different distribution for normal class data (0) and CAD class (1). The median age of CAD patients is much greater than

the median of healthy (normal) people, as are the values of quartiles 1 and 3. This means that the age feature is suspected to be a feature that influences whether someone has CAD or not. The Height, Weight, and BMI features have almost similar distributions between the normal class and CAD patients. In the Age feature, there are 2 outliers in class 0 data, this means that there are people who are old but do not have CAD.

From Figure 3, Table 2, and Table 3, it can be seen that the FBS (Fasting Blood Sugar), TG (Triglyceride), BUN (Blood Urea Nitrogen), and ESR (Erythrocyte Sedimentation Rate) features have higher median values for data class 1 (CAD patients) compared to data class 0 (normal). Meanwhile, CAD patients' HDL (High-Density Lipoprotein) and HB (Hemoglobin) features have smaller medians. Meanwhile, the LDL (Low-Density Lipoprotein) feature has almost the same median value for the normal and CAD data classes. The FBS, CR, and TG features of CAD patients have much larger interquartile range values than normal people. This shows that CAD patients' FBS, CR, and TG values are more spread out. Meanwhile, the interquartile range value of the HDL feature of CAD patients is smaller than that of normal people. This means that the HDL value of CAD patients is more clustered. All features in both class 0 and class 1 have outliers. This outlier can occur because of a measurement error or because there is outlier data.

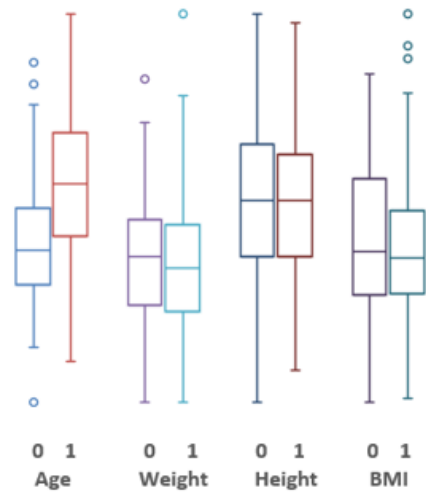


Figure 2. Box plot of demographic category features data

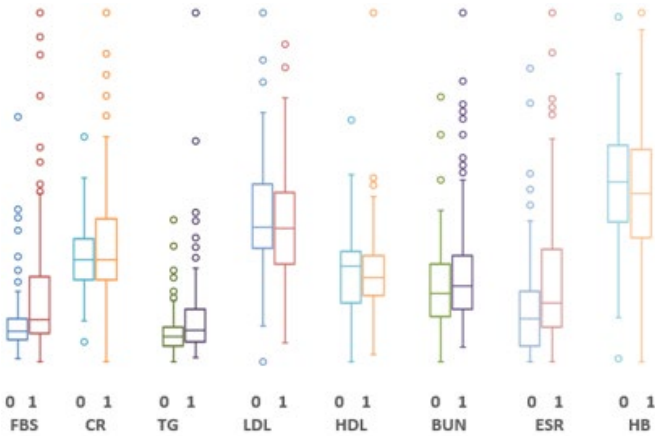


Figure 3. Box plot of laboratory and echo category features

Figure 4, Table 4, and Table 5 show that the K (potassium)

and Neut (neutrophil) features of CAD patients have higher median values than normal people. Meanwhile, CAD patients' Lymph (Lymphocyte) and EF-TTE (ejection fraction) features have smaller median values than normal people. The Na(sodium), WBC (white blood cell), and PLT (platelet) features have almost the same median values for data classes 0 and class 1. The Na feature for class 0 (normal) has a much smaller interquartile range, which means that the value of this feature for class 0 tends to be similar. There are many upper outliers in the WBC feature for class 1 data. This means that some CAD patients have a much larger number of white blood cells than their group.

Figure 5 and Table 6 show that the median BP (Blood Pressure) feature of CAD patients is greater than that of normal people. There are outliers in both CAD patients and normal people. Meanwhile, the PR (Pulse Rate) feature of class 0 shows a strange shape, where the median position coincides with quartile 1 or quartile 3. To find out, it cannot be seen visually. After looking at the numerical calculation, it turns out that the median value = quartile 1, which means that many data have the same value. There are upper outliers in the BP and PR features, which means that some normal people and CAD patients have higher BP and PR feature values than their group.

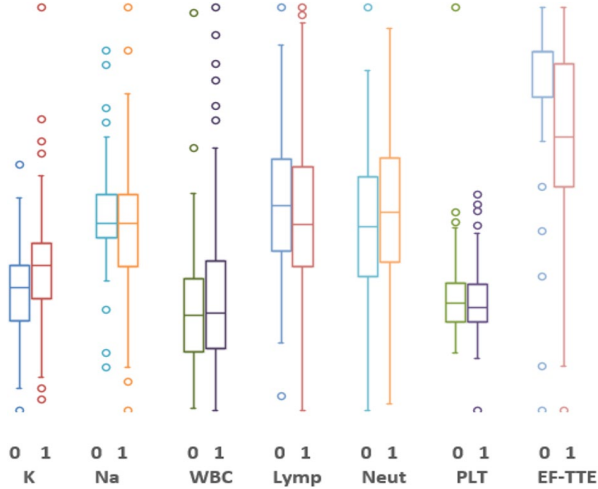


Figure 4. Box plot of laboratory and echo category features

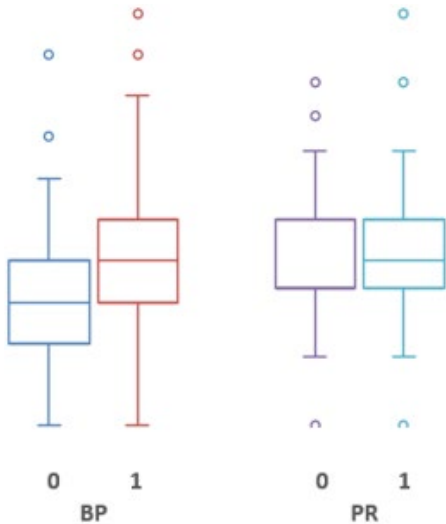


Figure 5. Box plot of symptoms and examination category features data

Table 1. Quartile 1, quartile 2, and quartile 3 values of age, weight, height, and BMI features for class 0 and 1 data

Quartiles	Age		Weight		Height		BMI	
	CAD 0	CAD 1	CAD 0	CAD 1	CAD 0	CAD 1	CAD 0	CAD 1
Q1	0.30	0.43	0.25	0.24	0.38	0.38	0.28	0.28
Q2	0.39	0.56	0.38	0.35	0.52	0.52	0.39	0.37
Q3	0.50	0.70	0.47	0.46	0.67	0.64	0.58	0.49

Table 2. Quartile 1, quartile 2, and quartile 3 values of FBS, CR, TG, and LDL features for class 0 and 1 data

Quartiles	FBS		CR		TG		LDL	
	CAD 0	CAD 1	CAD 0	CAD 1	CAD 0	CAD 1	CAD 0	CAD 1
Q1	0.07	0.08	0.24	0.24	0.05	0.06	0.33	0.28
Q2	0.09	0.12	0.29	0.29	0.07	0.09	0.39	0.38
Q3	0.12	0.24	0.35	0.41	0.10	0.15	0.51	0.48

Table 3. Quartile 1, quartile 2, and quartile 3 values of HDL, BUN, ESR, and HB features for class 0 and 1 data

Quartiles	HDL		BUN		ESR		HB	
	CAD 0	CAD 1	CAD 0	CAD 1	CAD 0	CAD 1	CAD 0	CAD 1
Q1	0.17	0.19	0.13	0.15	0.04	0.10	0.40	0.36
Q2	0.27	0.24	0.20	0.22	0.12	0.17	0.52	0.48
Q3	0.32	0.31	0.28	0.30	0.20	0.32	0.62	0.61

Table 4. Quartile 1, quartile 2, and quartile 3 values of K, Na, WBC, and Lymph features for class 0 and 1 data

Quartiles	K		Na		WBC		Lymph	
	CAD 0	CAD 1	CAD 0	CAD 1	CAD 0	CAD 1	CAD 0	CAD 1
Q1	0.22	0.28	0.43	0.36	0.15	0.15	0.40	0.36
Q2	0.31	0.36	0.46	0.46	0.24	0.24	0.51	0.46
Q3	0.36	0.42	0.54	0.54	0.33	0.37	0.62	0.60

Table 5. Quartile 1, quartile 2, and quartile 3 values of Neut, PLT, and EF-TTE features for class 0 and 1 data

Quartiles	Neut		PLT		EF-TTE	
	CAD 0	CAD 1	CAD 0	CAD 1	CAD 0	CAD 1
Q1	0.33	0.37	0.22	0.22	0.78	0.56
Q2	0.46	0.49	0.27	0.26	0.89	0.68
Q3	0.58	0.63	0.32	0.31	0.89	0.86

Table 6. Quartile 1, quartile 2, and quartile 3 values of BP and PR features for class 0 and 1 data

Quartiles	BP		PR	
	CAD 0	CAD 1	CAD 0	CAD 1
Q1	0,20	0,30	0,33	0,33
Q2	0,30	0,40	0,33	0,40
Q3	0,40	0,50	0,50	0,50

3.2 Classification

Table 7 shows the classification results on raw data. The best results are obtained when using demographic features, both in terms of accuracy, F1 score, and AUC. If normalization is first performed on features that have values greater than 1, better results are obtained, as seen in Table 8. The accuracy value changes from 0.77 to 0.87, the F1 score changes from 0.66 to 0.83, and the AUC value changes from 0.83 to 0.92.

Based on the accuracy value, the features that produce the best value are the symptom and examination features. While from the F1 score and AUC values, the best features are all features.

Different results are obtained for each evaluation metric if the outlier values are removed from the features. As seen in Table 9, the accuracy value after outlier removal decreases. Meanwhile, the F1 score value varies; some increase and some decrease, and so does the AUC value. From the AUC-ROC

curve in Figure 6, it can be seen that the symptom and examination features and all features can separate very well between class 0 and class 1.

Table 7. Classification results using several feature groups without normalization

No	Feature	Accu	F1 Score	AUC	Sens	Spec
1	Dem	0.77	0.66	0.83	0.89	0.4
2	ECG	0.76	0.43	0.78	1	0
3	Symp and exam	0.75	0.66	0.79	0.85	0.47
4	Lab and echo	0.71	0.50	0.55	0.89	0.13
5	All feature	0.75	0.43	0.5	1.0	0

Note: Dem: Demographic, ECG: Electrocardiogram, Symp and exam: Symptoms and examination, Lab and echo: Laboratory and echo, Accu: Accuracy, Sens: Sensitivity, Spec: Specificity.

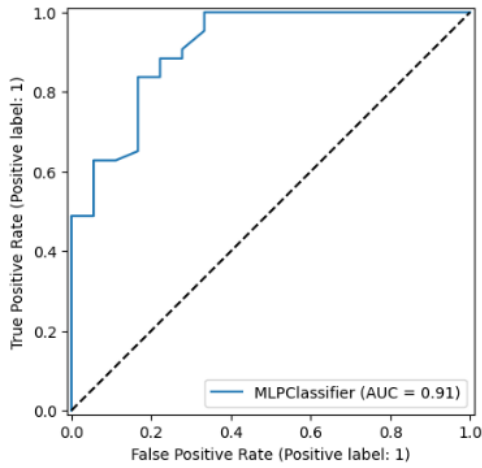
Table 8 Classification results using several feature groups with normalization

No	Feature	Accu	F1 Score	AUC	Sens	Spec
1	Dem	0.8	0.75	0.79	0.91	0.56
2	ECG	0.76	0.43	0.78	1	0
3	Symp and exam	0.87	0.83	0.91	0.95	0.67

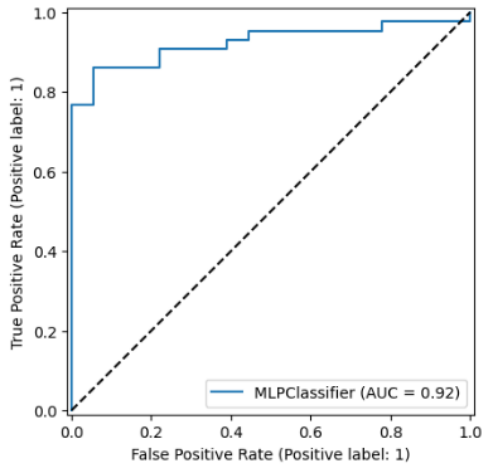
4	Lab and echo	0.71	0.67	0.69	0.72	0.67
5	All feature	0.85	0.83	0.92	0.88	0.78

Table 9. Classification results using several feature groups with outlier removal

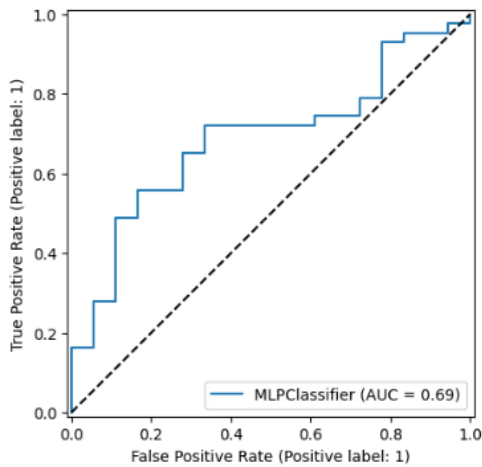
No	Feature	Accu	F1 Score	AUC	Sens	Spec
1	Dem	0.7996	0.76	0.79	0.897	0.6
2	Symp and exam	0.82	0.76	0.76	0.88	0.64
3	Lab and echo	0.69	0.64	0.73	0.72	0.6



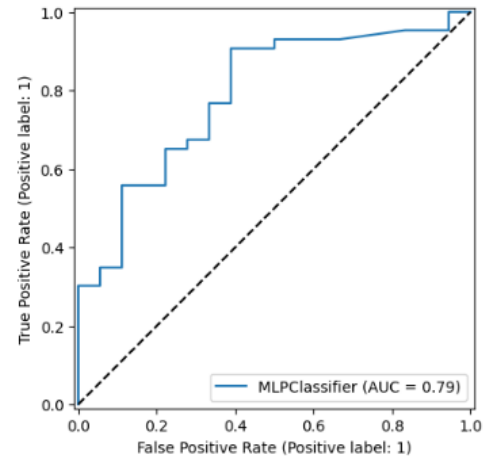
(a)



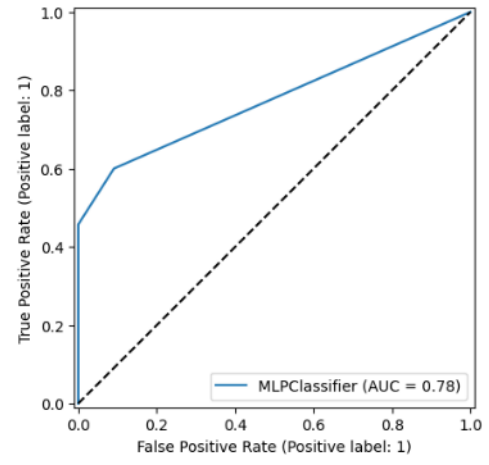
(b)



(c)



(d)

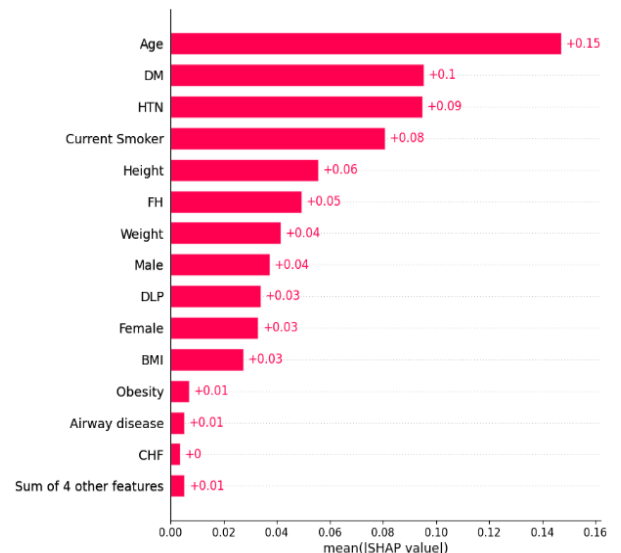


(e)

Figure 6. AUC-ROC curve of CAD classification problem using features (a) Symptoms and examination (b) All feature (c) Laboratory and echo (d) Demographic (e) ECG

3.3 Feature important

The classification results, as seen in Table 7, Table 8, and Table 9, only show the overall system performance. The SHAP value needs to be calculated to determine which features are most influential in determining the data class.



(a)

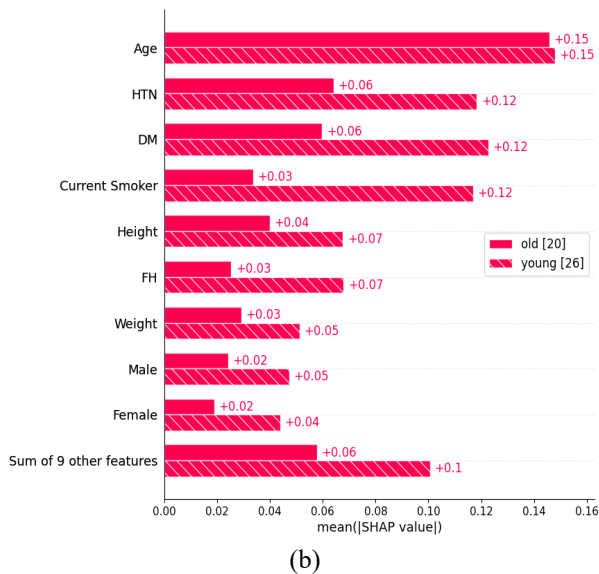


Figure 7. Average SHAP values of demographic features (a) age feature is not grouped (b) age feature is grouped into old (age > 58 years) and young (age ≤ 58 years)

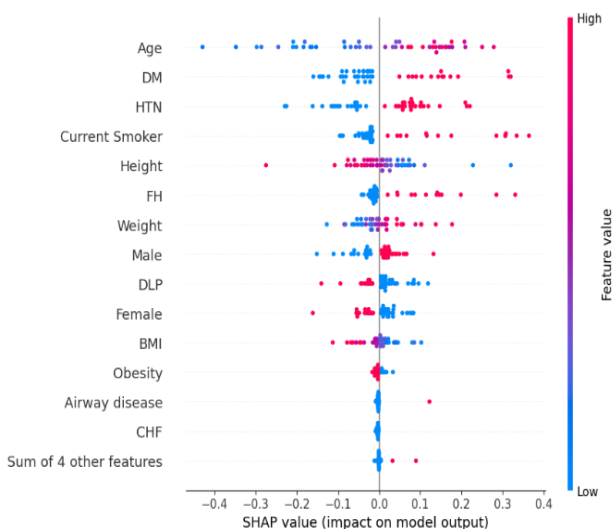


Figure 8. Beeswarm SHAP plot of demographic features

In Figure 7(a), age is the most influential demographic feature in determining the data class, followed by HTN (hypertension), DM (diabetes mellitus), and active smokers (current smokers).

Figure 7(a) shows that the influence of gender (male/female) is insignificant in influencing the data class. If the age feature is grouped into old (age > 58 years) and young (age ≤ 58 years), the average SHAP value graph is obtained as in Figure 7(b). The solid red bar chart indicates patients over 58 years old, and the red and white shaded color bar chart indicates patients ≤ 58 years old. Of the 46 testing data, 20 people are old, and 26 are young. From this age division, old or young does not have much effect on the data class. In the HTN (hypertension), DM (diabetes mellitus), and active smoker (current smoker) features, the young age group has more influence in determining whether someone has CAD or not.

The SHAP value of each testing data can be seen in Figure 8, a point represents testing data, and the color indicates the value of the feature. Low feature values are indicated by the

color blue; the thicker the blue color, the smaller the feature value; conversely, the red color indicates a high feature value; the thicker the red color, the greater the feature value. The point's color changes gradationally for features with continuous values such as age, weight, height, and BMI. Meanwhile, features that have Boolean values (0 or 1), such as DM, HTN, current smoker, FH (family history), gender (male and female), DLP (Dyslipidemia), Obesity, Airway Disease, Thyroid Disease, and CHF (congestive heart failure) the dot color is only dark red (if the feature has a value of 1) or dark blue (if the feature has a value of 0).

The age feature shows that the higher the age, the more positive the SHAP value, meaning that high age has a positive effect on the model output, or the higher the age, the greater the possibility of someone suffering from CAD. For the DM, HTN, and current smoker features, it can be seen that the red dot is on the right, and the blue dot is on the left. This means that patients with an attribute value of 1 (DM, HTN, current smoker is yes) have a high probability of suffering from CAD.

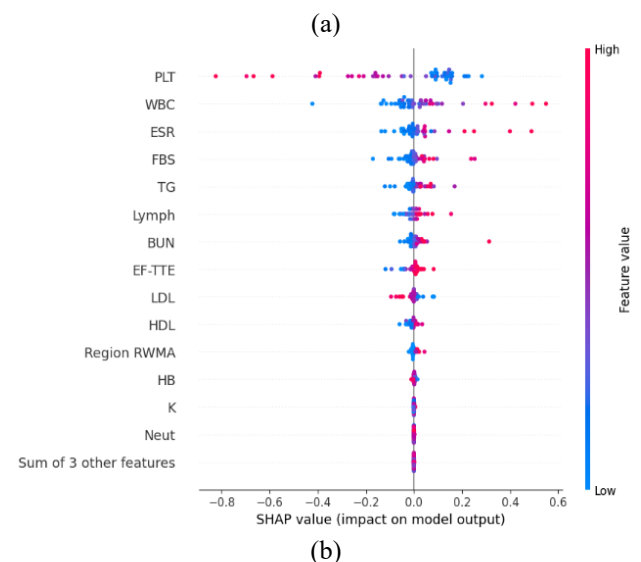
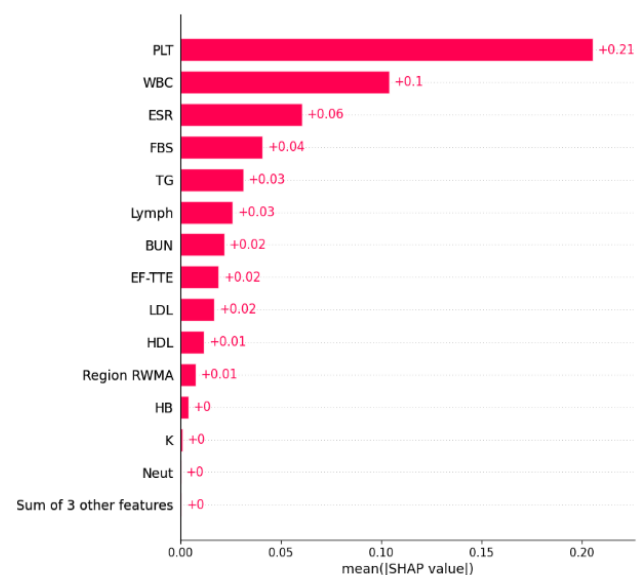


Figure 9. (a) Average SHAP values of laboratory and echo features, (b) SHAP beeswarm plot of laboratory and echo features

Figure 9(a) shows that the PLT (platelet) feature is the most influential feature of the Laboratory and echo category

features in determining the data class. From Figure 9(b), it can be seen that low platelet values positively affect the output, or low platelet values indicate that someone is suffering from CAD. The next most influential feature of the laboratory and echo group is the WBC (white blood cell) feature, or the number of white blood cells per milliliter, ranging from 3700 -18,000. From Figure 9(b), it can be seen that high white blood cell counts positively affect data class 1.

Figure 10(a) shows that the 5 features that have the most influence on determining the data class from the symptom and examination feature category are typical chest pain, followed by atypical, dyspnea, function class, and BP features. From Figure 10(b), it can be concluded that a high typical chest pain value (valued 1 or typical chest pain = yes) positively affects the output. In other words, patients with certain chest pain symptoms indicate that they have CAD. Meanwhile, the opposite applies to the Atypical and Dyspnea attributes; attributes with a 0 (no) positively affect the output.

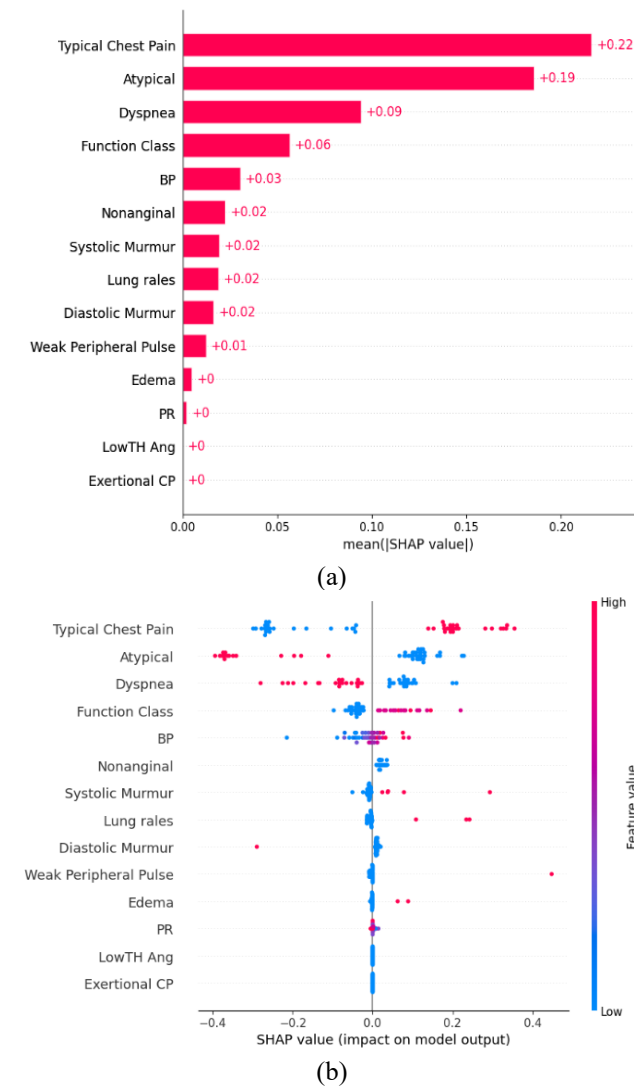


Figure 10. (a) Average SHAP values of symptom and examination features, (b) SHAP beeswarm plot of symptom and examination features

3.4 Local analysis

Local analysis aims to explain a particular data point. Test data #10, #20, #30, and #51 are taken.

Figure 11 shows LIME's explanation of testing data #10.

According to LIME, the testing data class is the Cath class, with a probability of 0.58. The features that influence this testing data to enter the CATH class are BMI, which has a value of 24.49; CHF, which has a value of 0; CRF, which has a value of 0; and Thyroid disease, which has a value of 0. Age, height, gender, and current smokers negatively influence Cath.

Figure 12 shows the LIME results of the 51st testing data in the Normal class. The features that influence this testing data to enter the normal class are BMI, which is 26.26, Current smoking, which is 0 (does not smoke), and Male gender. The Age, HTN, height, DM, and CHF features support the Cath class data entry.

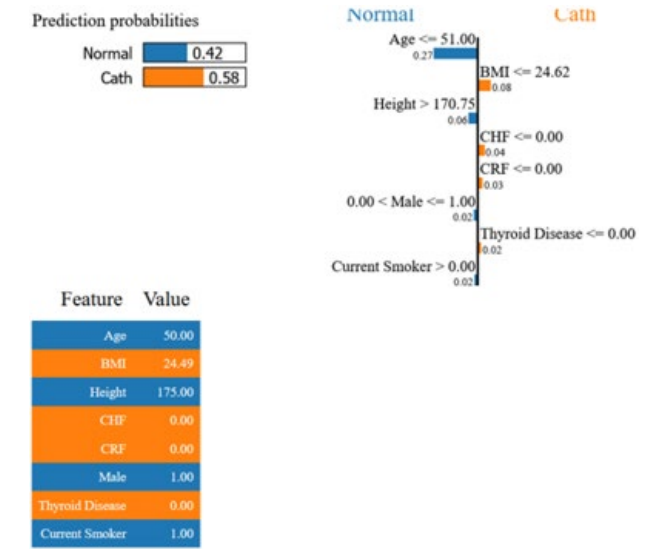


Figure 11. LIME's tabular explanation for observation #10 (CATH)

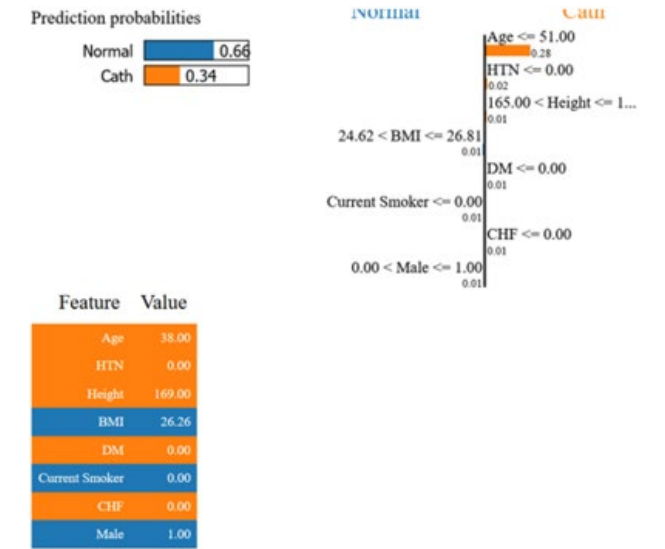


Figure 12. LIME's tabular explanation for observation #51 (Normal)

To see the consistency of LIME prediction results on certain testing data, LIME was run 10 times for testing data #10, #20, #30, and #51. Table 10 shows the results of running LIME for data #10 10 times. As did the influential features, the predicted probability values for the normal and Cath classes varied in each LIME run. Only in the 6th, 7th, and 8th runs were the predicted probability values for the class produced the same value, but the influential features were different. Although the

predicted probability values for the normal and Cath classes differed, testing data #10 was still classified as the Cath class (class 1). From the 10 LIME runs, the BMI feature appeared the most, influencing data #10 to enter the Cath class.

Table 11 shows the results of running LIME 10 times on testing data #51. In addition to the different prediction probability values for each run, different classification results were also obtained on the 8th and 9th runs. Testing data #51 is recognized as the Cath class, while in other runs, it is

recognized as the normal class. The most influential features in determining the class for testing data #51 are also more varied.

In Table 12 and Table 13, it can be seen that the features that influence the test data to enter a class vary for each LIME run. The changing results for each run of LIME on a particular testing data set are caused by randomly generating new datasets around the testing data to be explained.

Table 10. Results of 10 runs of LIME from testing data #10

No	Prediction Probabilities		The Two Most Influential Features
	Normal	Cath	
1	0.36	0.64	BMI = 24.49, Airways diseases = 0
2	0.45	0.55	Female = 0 (male), CVA = 0
3	0.44	0.56	BMI = 24.49, Weight = 75
4	0.27	0.73	Height = 175, Weight = 75
5	0.47	0.53	BMI = 24.49, Weight = 75
6	0.42	0.58	BMI = 24.49, CHF = 0
7	0.42	0.58	BMI = 24.49, CVA = 0
8	0.42	0.58	CVA = 0, Ex-smoker = 0
9	0.37	0.63	BMI = 24.49, CVA = 0
10	0.31	0.69	Curent smoker = 1, Male = 1

Table 11. Results of 10 runs of LIME from testing data #51

No	Prediction Probabilities		The Two Most Influential Features
	Normal	Cath	
1	0.66	0.34	BMI = 26.26, Current smoker = 1
2	0.57	0.43	BMI = 26.26, Weight = 75
3	0.53	0.47	Obesity = 0, CRF = 0
4	0.63	0.37	Female = 0 (male), Thyroid diseases = 0
5	0.65	0.35	CRF = 0, DLP = 0
6	0.58	0.42	CHF = 0, CVA = 0
7	0.69	0.31	CVA = 0, CHF = 0
8	0.29	0.71	Thyroid diseases = 0, Airway diseases = 0
9	0.43	0.57	Age = 38, Height = 169
10	0.61	0.39	CHF = 0, Female = 0 (male)

Table 12. Results of 10 runs of LIME from testing data #20

No	Prediction Probabilities		The Two Most Influential Features
	Normal	Cath	
1	0.49	0.51	WBC = 4100, CRF = 0
2	0.38	0.62	PLT = 161, Q Wave = 0
3	0.29	0.71	PLT = 161
4	0.27	0.73	CHF = 0, CVA = 0
5	0.28	0.72	Age = 65, Weight = 78
6	0.34	0.66	CHF = 0, CVA = 0
7	0.3	0.7	CHF = 0
8	0.3	0.7	Nonanginal = 0, FH = 0
9	0.3	0.7	Age = 65, Weight = 78
10	0.3	0.7	Age = 65, Weight = 78

Table 13. Results of 10 runs of LIME from testing data #30

No	Prediction Probabilities		The Two Most Influential Features
	Normal	Cath	
1	0.26	0.74	Nonanginal = 0, Q Wave = 0
2	0.29	0.71	Age = 80, Weight = 60
3	0.31	0.69	Age = 80, Weight = 60
4	0.32	0.68	St Depression = 0, Atypical = 0
5	0.3	0.7	LowTHAng = 0, CHF = 0
6	0.32	0.68	CHF = 0, Weak Pheriperal Pulse = 0
7	0.27	0.73	HDL = 49, BMI = 19.82
8	0.31	0.69	Age = 80, Weight = 60
9	0.3	0.7	Systolic Murmur = 0, FH = 0
10	0.29	0.71	Age = 80, Weight = 60

3.5 Discussion

Applying EML to classify CAD-RADS scores based on heart disease risk factors has been successfully carried out. The classification results show quite good system performance, both measured by accuracy, F1 score, AUC, sensitivity, and Specificity. The most influential features in determining the data class can be determined using the SHAP value. This result is better than the traditional machine learning method (black box model), which can only determine model performance but cannot determine which features have the most influence on determining data class. However, local analysis to determine the most influential features of a particular data set is not very good. This is indicated by the inconsistent LIME results between the results of running the program. This will be a challenge if this method is applied to real-world settings.

4. CONCLUSION

This paper has discussed the relationship between heart disease risk factors and CAD-RADS scores. It starts by performing EDA on the feature categories. The EDA results provide an initial picture of the features that influence the classification. For classification purposes, the highest performance is achieved when using all features. The features that most influence class determination can be determined using SHAP. LIME can be used to create a model that can be interpreted from the black box model on a particular testing data. From the results of the local analysis, it can be seen that the results obtained vary for each run of LIME. Both from the results of the probability of predicting normal and CAD classes, as well as from the most influential features. This occurs because of the random process of generating data around the data to be explained. Various sampling techniques can be studied for further research that minimizes variation between LIME runs.

REFERENCES

- [1] Cury, R.C., Leipsic, J., Abbara, S., Achenbach, S., et al. (2022). CAD-RADS™ 2.0–2022 coronary artery disease-reporting and data system: An expert consensus document of the Society of Cardiovascular Computed Tomography (SCCT), the American College of Cardiology (ACC), the American College of Radiology (ACR), and the North American Society of Cardiovascular Imaging (NASCI). *Cardiovascular Imaging*, 15(11): 1974-2001. <https://doi.org/10.1148/ryct.220183>
- [2] Bampi, A.B.A., Rochitte, C.E., Favarato, D., Lemos, P.A., da Luz, P.L. (2009). Comparison of non-invasive methods for the detection of coronary atherosclerosis. *Clinics*, 64(7): 675-682. <https://doi.org/10.1590/S1807-59322009000700012>
- [3] Tridandapani, S., Banait-Deshmane, S., Aziz, M.U., Bhatti, P., Singh, S.P. (2021). Coronary computed tomographic angiography: A review of the techniques, protocols, pitfalls, and radiation dose. *Journal of Medical Imaging and Radiation Sciences*, 52(3): S1-S11. <https://doi.org/10.1016/j.jmir.2021.08.014>
- [4] Abdelrahman, K.M., Chen, M.Y., Dey, A.K., Virmani, R., et al. (2020). Coronary computed tomography angiography from clinical uses to emerging technologies: JACC state-of-the-art review. *Journal of the American College of Cardiology*, 76(10): 1226-1243. <https://doi.org/10.1016/j.jacc.2020.06.076>
- [5] Kheiri, B., Simpson, T.F., Osman, M., German, D.M., Fuss, C.S., Ferencik, M. (2022). Computed tomography vs invasive coronary angiography in patients with suspected coronary artery disease: A meta-analysis. *Cardiovascular Imaging*, 15(12): 2147-2149. <https://doi.org/10.1016/J.JCMG.2022.06.003>
- [6] Xie, J.X., Cury, R.C., Leipsic, J., Crim, M.T., et al. (2018). The coronary artery disease-reporting and data system (CAD-RADS) prognostic and clinical implications associated with standardized coronary computed tomography angiography reporting. *JACC: Cardiovascular Imaging*, 11(1): 78-89. <https://doi.org/10.1016/j.jcmg.2017.08.026>
- [7] Kumar, P., Bhatia, M. (2021). Coronary artery disease reporting and data system: A comprehensive review. *Journal of Cardiovascular Imaging*, 30(1): 1. <https://doi.org/10.4250/jcvi.2020.0195>
- [8] Denzinger, F., Wels, M., Breininger, K., Gülsün, M.A., et al. (2020). Automatic CAD-RADS scoring using deep learning. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru*, pp. 45-54. https://doi.org/10.1007/978-3-030-59725-2_5
- [9] Wahab Sait, A.R., Dutta, A.K. (2023). Developing a deep-learning-based coronary artery disease detection technique using computer tomography images. *Diagnostics*, 13(7): 1312. <https://doi.org/10.3390/diagnostics13071312>
- [10] Huang, Z., Xiao, J., Wang, X., Li, Z., et al. (2023). Clinical evaluation of the automatic coronary artery disease reporting and data system (CAD-RADS) in coronary computed tomography angiography using convolutional neural networks. *Academic Radiology*, 30(4): 698-706. <https://doi.org/10.1016/j.acra.2022.05.015>
- [11] Huang, Z., Yang, Y., Wang, Z., Hu, Y., et al. (2023). Comparison of prognostic value between CAD-RADS 1.0 and CAD-RADS 2.0 evaluated by convolutional neural networks based CCTA. *Heliyon*, 9(5): e15988. <https://doi.org/10.1016/j.heliyon.2023.e15988>
- [12] Dai, Y., Ouyang, C., Luo, G., Cao, Y., Peng, J., Gao, A., Zhou, H. (2023). Risk factors for high CAD-RADS scoring in CAD patients revealed by machine learning methods: A retrospective study. *PeerJ*, 11: e15797. <https://doi.org/10.7717/PEERJ.15797>
- [13] Alizadehsani, R., Habibi, J., Hosseini, M.J., Mashayekhi, H., et al. (2013). A data mining approach for diagnosis of coronary artery disease. *Computer Methods and Programs in Biomedicine*, 111(1): 52-61. <https://doi.org/10.1016/j.cmpb.2013.03.004>
- [14] Baig, N., Abba, S.I., Usman, J., Muhammad, I., Abdulazeez, I., Usman, A.G., Aljundi, I.H. (2024). Bio-inspired MXene membranes for enhanced separation and anti-fouling in oil-in-water emulsions: SHAP explainability ML. *Cleaner Water*, 2: 100041. <https://doi.org/10.1016/j.clwat.2024.100041>
- [15] Zhang, X., Dimitrov, N. (2024). Variable importance analysis of wind turbine extreme responses with Shapley value explanation. *Renewable Energy*, 232: 121049. <https://doi.org/10.1016/j.renene.2024.121049>

- [16] Borgonovo, E., Plischke, E., Rabitti, G. (2024). The many Shapley values for explainable artificial intelligence: A sensitivity analysis perspective. *European Journal of Operational Research*, 318(3): 911-926. <https://doi.org/10.1016/j.ejor.2024.06.023>
- [17] Anjum, M., Khan, K., Ahmad, W., Ahmad, A., Amin, M.N., Nafees, A. (2022). New SHapley additive exPlanations (SHAP) approach to evaluate the raw materials interactions of steel-fiber-reinforced concrete. *Materials*, 15(18): 6261. <https://doi.org/10.3390/ma15186261>
- [18] Lu, Y., Fan, X., Zhang, Y., Wang, Y., Jiang, X. (2023). Machine learning models using SHapley Additive exPlanation for fire risk assessment mode and effects analysis of stadiums. *Sensors*, 23(4): 2151. <https://doi.org/10.3390/s23042151>
- [19] Tiwari, R.S. (2024). Hate speech detection using LSTM and explanation by LIME (local interpretable model-agnostic explanations). In *Computational Intelligence Methods for Sentiment Analysis in Natural Language Processing Applications*, pp. 93-110. <https://doi.org/10.1016/B978-0-443-22009-8.00005-7>
- [20] Rodríguez-Pérez, R., Bajorath, J. (2019). Interpretation of compound activity predictions from complex machine learning models using local approximations and Shapley values. *Journal of medicinal chemistry*, 63(16): 8761-8777. <https://doi.org/10.1021/acs.jmedchem.9b01101>