

Article

Implementing an Ensemble Learning Model with Feature Selection to Predict Mortality among Patients Who Underwent Three-Vessel Percutaneous Coronary Intervention

Yen-Chun Huang ^{1,2,†} , Kuan-Yu Chen ^{2,3,†}, Shao-Jung Li ^{4,5,6,7}, Chih-Kuang Liu ^{1,2,8}, Yang-Chao Lin ^{2,9}  and Mingchih Chen ^{1,2,*} 

- ¹ Artificial Intelligence Development Center, Fu Jen Catholic University, No. 510, Zhongzheng Rd., Xinzhuang Dist., New Taipei City 242, Taiwan
 - ² Graduate Institute of Business Administration, College of Management, Fu Jen Catholic University, No. 510, Zhongzheng Rd., Xinzhuang Dist., New Taipei City 242, Taiwan
 - ³ Division of Cardiology, Taipei City Hospital, Zhongxing Branch, Taipei 106, Taiwan
 - ⁴ Cardiovascular Research Center, Wan Fang Hospital, Taipei Medical University, Taipei 242, Taiwan
 - ⁵ Taipei Heart Institute, Taipei Medical University, Taipei 242, Taiwan
 - ⁶ Department of Surgery, School of Medicine, College of Medicine, Taipei Medical University, Taipei 242, Taiwan
 - ⁷ Division of Cardiovascular Surgery, Department of Surgery, Wan Fang Hospital, Taipei Medical University, Taipei 242, Taiwan
 - ⁸ Department of Urology, Fu Jen Catholic University Hospital, New Taipei City 243, Taiwan
 - ⁹ Department of Gastroenterology and Hepatology, Fu Jen Catholic University Hospital, New Taipei City 243, Taiwan
- * Correspondence: 081438@mail.fju.edu.tw
- † These authors contributed equally to this work.



Citation: Huang, Y.-C.; Chen, K.-Y.; Li, S.-J.; Liu, C.-K.; Lin, Y.-C.; Chen, M. Implementing an Ensemble Learning Model with Feature Selection to Predict Mortality among Patients Who Underwent Three-Vessel Percutaneous Coronary Intervention. *Appl. Sci.* **2022**, *12*, 8135. <https://doi.org/10.3390/app12168135>

Academic Editors: Mohamed Medhat Gaber and Mohammed Abdelsamea

Received: 30 June 2022

Accepted: 8 August 2022

Published: 14 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Coronary artery disease (CAD) is a common major disease. Revascularization with percutaneous coronary intervention (PCI) or coronary artery bypass graft (CABG) could relieve symptoms and myocardial ischemia. As the treatment improves and evolves, the number of aged patients with complex diseases and multiple comorbidities gradually increases. Furthermore, in patients with multivessel disease, 3-vessel PCI may lead to a higher risk of complications during the procedure, leading to further ischemia and higher long-term mortality than PCI for one vessel or two vessels. Nevertheless, the risk factors for accurately predicting patient mortality after 3-vessel PCI are unclear. Thus, a new risk prediction model for primary PCI (PPCI) patients' needs to be established to help physicians and patients make decisions more quickly and accurately. This research aimed to construct a prediction model and find which risk factors will affect mortality in 3-vessel PPCI patients. This nationwide population-based cohort study crossed multiple hospitals and selected 3-vessel PPCI patients from January 2007 to December 2009. Then five different single machine learning methods were applied to select significant predictors and implement ensemble models to predict the mortality rate. Of the 2337 patients who underwent 3-vessel PPCI, a total of 1188 (50.83%) survived and 1149 (49.17%) died. Age, congestive heart failure (CHF), and chronic renal failure (CRF) are mortality's most important variables. When CRF patients accept 3-vessel PPCI at ages between 68–75, they will possibly have a 94% death rate; Furthermore, this study used the top 15 variables averaged by each machine learning method to make a prediction model, and the ensemble learning model can accurately predict the long-term survival of 3-vessel PPCI patients, the accurate predictions rate achieved in 88.7%. Prediction models can provide helpful information for the clinical physician and enhance clinical decision-making. Furthermore, it can help physicians quickly identify the risk features, design clinical trials, and allocate hospital resources effectively.

Keywords: percutaneous coronary intervention; National Health Insurance Research Database; mortality prediction; machine learning; ensemble learning model; feature selection

1. Introduction

Coronary artery disease (CAD) is a common global health problem and is the main cause of death [1,2]. The main purpose of using percutaneous coronary intervention (PCI) or coronary artery bypass graft (CABG) to treat patients with multivessel CAD is to relieve myocardial ischemia and other symptoms. PCI for multivessel diseases can result in inadequate revascularization or contrast-induced acute kidney injury. There is also a 50% probability of cardiac or noncardiac death associated with this procedure [3,4].

Compared to CABG surgery, PCI is relatively less invasive. PCI is increasingly being accepted as a treatment option with improved skills and instruments [5]. This procedure can be used to treat patients who refuse or cannot tolerate CABG and is estimated to gradually become the treatment of choice for CAD [6–8].

Although many studies mention PCI as the optimal treatment, patient comorbidities can affect morbidity and mortality [9–17]. When three vessels are treated simultaneously, the probability of complications, or even death, during or after surgery is higher than when one or two are treated [18]. Many risk factors for mortality after PCI have been identified [19,20], including clinical and anatomic variables. Clinical variables include age, sex, diabetes mellitus, chronic lung disease, prior myocardial infarction, impairment of left ventricular function, renal dysfunction, cardiogenic shock, left main coronary artery (LMCA), and age are the most important predictors of 1-year mortality postoperatively [21]. Thus, postoperative death remains a major issue for PCI.

Several studies have only used single machine learning methods to make the prediction model, such as focused on the degree of revascularization [22,23] or logistic regression (LGR) for predicting in-hospital mortality [24]. Nevertheless, some mortality prediction models after PCI have been constructed using the data of significantly older people or clinical factors [25–28]. No study has focused on mortality risk factors among patients who initially underwent 3-vessel PCI, and there are no machine learning prediction models.

The present study collected data from the Taiwan National Health Insurance Research Database (NHIRD), one of Asia's largest databases, and aimed to identify variables that affect the survival of patients undergoing 3-vessel primary PCI (PPCI). The patient's risk factors and baseline characteristics were collected the year before surgery. After identifying the optimal features, ensemble learning methods were used to construct a prediction model. Five machine learning methods were applied, and majority voting was used to develop the model. This prediction model can be a basis for clinical decision-making and healthcare management policies.

2. Materials and Methods

2.1. Data Source

Data from patients who underwent 3-vessel PCI for the first time were collected from the NHIRD. The NHI system was launched in March 1995, and approximately 23 million beneficiaries joined this program. The NHIRD collects patient demographics, treatments, medication, total expenditure, admissions, and outpatient medical records from every medical institution. The diagnosis code identifies a patient's history, as defined by the International Classification of Diseases, 9th Revision. Clinical Modifications ICD-9-CM and ICD-10-CM were added to the database in 2016. The Institutional Review Board of Fu Jen Catholic University Ethics Institutional Review Board in Taiwan approved the protocol of this study and waived the need for informed consent (IRB approval number: C108121).

2.2. Study Framework

The main goal of this research was to identify the factors affecting the mortality of patients undergoing 3-vessel primary PCI and construct a prediction machine learning model. For this aim, a three-stage framework was designed. In the first stage, data of patients who underwent 3-vessel PCI for the first time between 1 January 2007, and 31 December 2009, were extracted (operation codes 33078A and 33078B) from the NHIRD in Taiwan ($n = 2647$). The exclusion criteria were missing information ($n = 6$) and patients

who underwent PCI in 2002–2006 ($n = 304$). Finally, 2337 patients were enrolled in this study; 1188 (50.83%) survived, and 1149 (49.17%) died. The initial inpatient date of PCI was set as the index date. The final follow-up time was the day of death or follow-up until the end of the time in the database (31 December 2019). Death records were extracted from the National Death Registry. Figure 1 presents the overall conceptual framework.

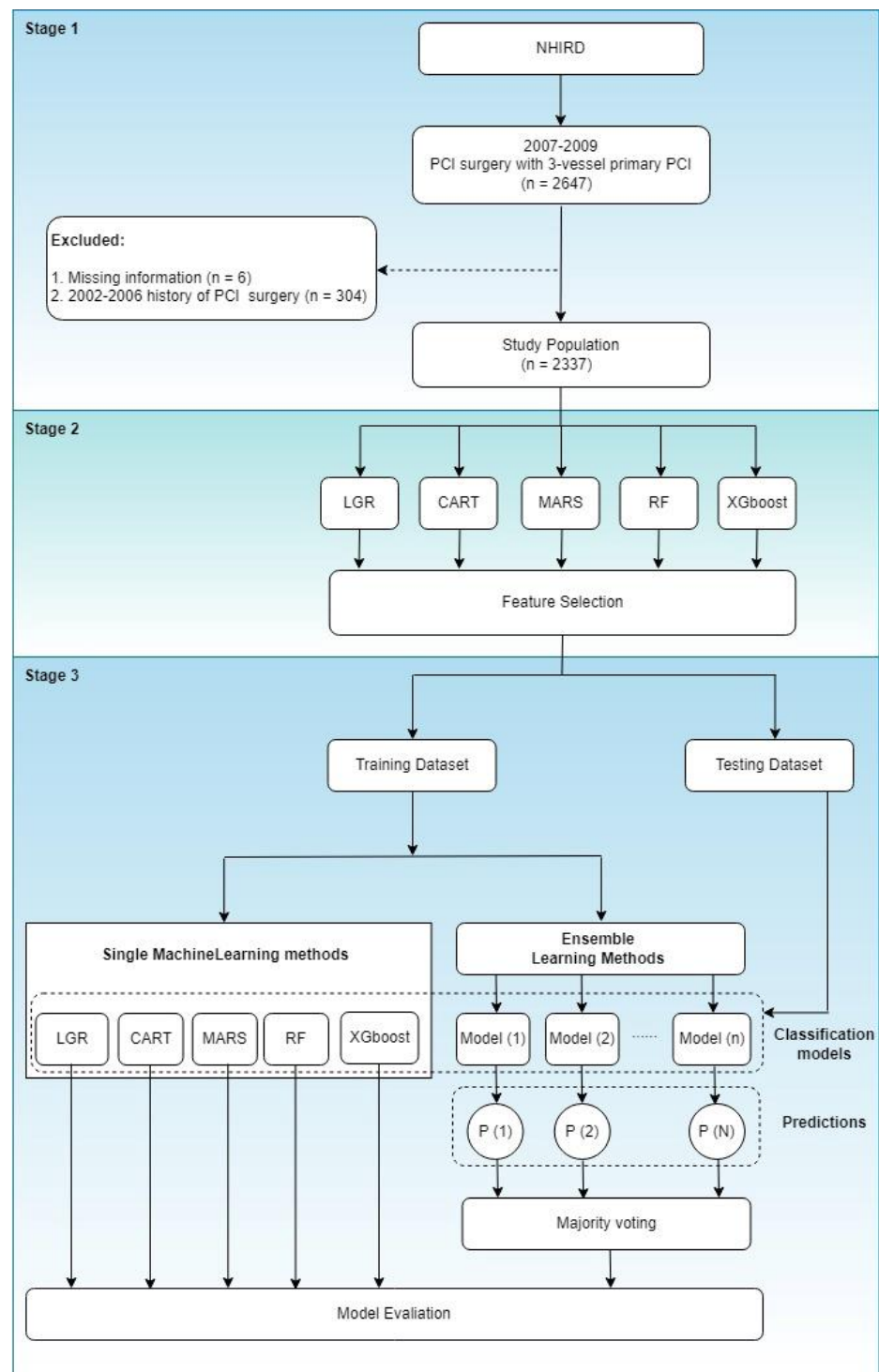


Figure 1. Processing framework of patients who underwent first-time PCI between 2007 to 2009. CART, classification and regression tree; LGR, logistic regression; MARS, multivariate adaptive regression splines; NHIRD, National Health Insurance Research Database; PCI, percutaneous coronary intervention; RF, random forests; XGboost, extreme gradient boosting.

After selecting the population (stage 2), we trained five different machine learning models using five cross-validations to find the optimal features. A fivefold cross-validation approach was applied to divide five equal subset folds in the training dataset randomly, and four subset folds were used to build the model as a training dataset. The remaining fold (1-fold) was used to validate the model. The five machine models included LGR, classification and regression tree (CART), multivariate adaptive regression splines (MARS), random forests (RF), and extreme gradient boosting (XGboost). In the final stage, the feature scores for each method were averaged, and single and ensemble prediction models were implemented; finally, the final prediction result was evaluated.

2.3. Risk Features

Overall, 42 risk factors that affected the survival rate in the patients were selected. The baseline variables in X1–X3 were sex (male/female), age, and Charlson Comorbidity Index (CCI). The X4–X7 variables included the preoperative 1-year history, including previous CABG surgery, the average length of ICU stays, average bags of blood transfused (94001C, 94002C, 4013C, 94015C, 94003C), and average days on mechanical ventilation (57001B, 57002B, 57003B). X8–X11 were surgical variables, and X12–X39 were all underlying diseases.

Underlying diseases were divided into the following: baseline characteristics, including hypertension (X12), hyperlipidemia (X13), hyperuricemia (X14), diabetes (X15), liver cirrhosis (X16), chronic obstructive pulmonary disease (COPD; X17), skin and bone diseases (X18), stroke (X19), gout (X20), biliary stones (X21), hepatitis B virus (HBV, X22), and hepatitis C virus (HCV, X23); heart-related characteristics, including atrial fibrillation (AF, X24), CHF (X25), peripheral vascular disease (PVD, X26), acute coronary syndrome (ACS, X27), malignant dysrhythmia (X28), and cardiogenic shock (X29); renal-related diseases, including kidney disease (X30; contain: glomerulonephritis, nephritis, acute renal failure, and chronic renal failure), chronic renal failure (CRF, X31), and acute kidney failure (AKF, X32); infection-related characteristics, including septicemia (X33), lower respiratory tract infections (Lower RTI, X34), and gastrointestinal infections (GTIs, X35); and bleeding-related characteristics intracranial bleeding (X36), transient ischemic attack (TIA, X37), gastrointestinal bleeding (X38), and major bleeding (X39).

Hospital variables included the following: hospital status (medical/nonmedical center) (X40), hospital ownership (public/private hospital) (X41), and hospital area (central/northern/southern/eastern) (X42).

2.4. Outcome Definition

According to a previous study, medical records of up to 1 year were useful predictors of mortality [29]. Therefore, data from inpatient and outpatient records in the previous year and the variables of current surgery were included as predictive variables. The primary outcome of death was collected from the Taiwan Cancer Registry files.

3. Feature Selection and Methodology

The medical information for patients was extensive and poorly organized. Checking past medical records before making clinical decisions is time-consuming. When estimating risk factors, essential variable selection is a crucial preprocessing step. Variable selection excludes extraneous risk factors while focusing on strongly correlated features without losing information [29–31]. The prediction model was designed to identify the risk features affecting mortality, and the optimal features were used to construct the prediction model for 3-vessel PCI patients. Machine learning involves using data to improve learning automatically. The most commonly used method is LGR. In the present study, we applied five different single machine learning methods, extracted the advantage from each model, and used majority voting to ensure an optimal ensemble learning model. The predicted accuracy of the above methods was then evaluated and used to calculate accuracy, kappa, sensitivity, specificity, AUC, and F1_Score. All machine learning prediction methods were implemented in the R studio software with version 3.6.2. The “glm” package was used for

LGR, “rpart” was used for CART, “earth” was used for MARS, “randomForest” was used for RF, and “xgboost” was used for XGBoost. The parameter is shown in supplementary (Table S1).

3.1. Feature Selection

Prompt decision-making in a limited time is challenging for physicians. This makes feature selection significantly important. Feature selection involves preprocessing to reduce computation time, improve prediction performance, and identify the optimal features for overcoming problems.

To enhance the accuracy of prediction, multiple collinearities and low correlation features between dependent variables should be eliminated, then the rate can be improved [32]. Thus, we take advantage of each machine learning algorithm to find which risk features correlate highly with PCI death.

3.2. LGR

LGR is a classical statistical classification model that evolved from linear regression. LGR can be used to perform binary classification tasks by predicting the probabilities of outcomes. The model offers only two possible outcomes: Y ; e.g., yes (1) or no (0). LGR can be used to analyze the correlation between one or more independent variables. The basic Equation (1) for LGR is as follows:

$$\log\left(\frac{b}{1-b}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 \dots \beta_i x_i, \quad (1)$$

where β_0 is the intercept term or error, β_1 is the coefficient of the independent variable (x_1), and the output number b is the probability between 0 and 1.

3.3. CART

The decision tree is a flexibility classification model, and the rule is defined as if/else instruction [33]. This structure of decision-tree formation using a classifier method involves the creation of an observation model and classifying input data by the tree from the node to the target or node to the node. The branches are the link between each node [34].

3.4. RF

RF comprises many decision trees, and the resulting forest is trained by bagging or bootstrap aggregation. It can improve the accuracy of machine learning algorithms and prevent the overfitting of the dataset. The final result is obtained by combining each decision tree's mode or weighted average to make predictions. With the use of out-of-bag (OOB) error, it is possible to evaluate the constructed model's performance and determine the input factors' importance [35,36]. Compared with ANN, RF has the advantage of allowing the improvement of predictive performance and easier understanding [35,37].

3.5. MARS

MARS is a nonparametric statistical method proposed by Friedman et al. in 1991 [38]. It can automatically create standard models and handle high-dimensional data [39]. MARS proceeds in a forward-backward pattern. Typically, pair basis functions are added to the model in the forward aspect until the maximum value is reached. In the backward aspect, an overfit model with being built, and the least efficient terms will be pruned until the best sub-model is found [39,40]. The MARS model is presented in Equation (2).

$$f(X) = \beta_0 + \sum_{i=1}^I \beta_i * \lambda_i * (Y), \quad (2)$$

where $\lambda_i * (Y)$ can be one or more spline functions, and β (constant) is the coefficient calculated using the least squares estimation method to estimate the coefficients.

3.6. XGboost

XGboost is an ensemble machine learning algorithm based on trees and uses a gradient boosting technique to significantly reduce learning time and regulation function to prevent overfitting [35]. Boosting can group weak learners into a set and predict accuracy. When the gradient descents, it will weigh the error of weak learning prediction models and present them sequentially in the next learned model to minimize loss function fitting the residuals and construct a strong prediction model. The loss function derives from Taylor's expansion.

3.7. Ensemble Modeling Methods

No single prediction model can achieve perfect accuracy because, under different data and assumptions, different algorithms have different inductive biases. Therefore, ensemble learning takes advantage of each classifier to improve accuracy and combines them to reach the final result [41,42]. The main concept of ensemble methodology is constructing multiple individual classifiers according to each learner's performance and stability, then combining them to obtain better performance in the single classifier [43,44]. Ensemble modeling has been successfully applied to various decision-making problems, such as feature selection, classification, and prediction [41]. It can reduce the estimated error variance and avoid overfitting problems [42,43,45].

3.8. Performance Metrics

The performance metrics are common to use in model evaluation to evaluate the performance of each prediction model, and it is calculated by confusion matrix value: true positive (TP), true negative (TN), false positive (FP), and false negative (FN). TP and TN are the cases where the actual value is correctly predicted. FP and FN are observed when the case predicted the actual value was incorrect. The equations for the metrics of accuracy, kappa, sensitivity, specificity, and F1 score are defined in Equations (3)–(7), respectively:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} * 100 \quad (3)$$

$$y = \frac{\left[(\text{TP} + \text{FP}) \times \left(\frac{\text{TP}}{\text{Accuracy}} \right) \right] + [(\text{FN} + \text{TN}) \times (\text{Accuracy})]}{\text{Accuracy}} \quad (4)$$

$$\text{Kappa} = \frac{(y - \text{Accuracy})}{(100\% - y)}$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (5)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (6)$$

$$\text{F1 - score} = 2 \times \frac{\text{Sensitivity} \times \text{Specificity}}{\text{Sensitivity} + \text{Specificity}} \quad (7)$$

Hosmer et al. categorized AUC values as follows: $\text{AUC} \geq 0.9$, outstanding discrimination; $0.8 \leq \text{AUC} < 0.9$, good discrimination; $0.7 \leq \text{AUC} < 0.8$, acceptable/fair discrimination; $0.6 \leq \text{AUC} < 0.7$, poor discrimination; and $\text{AUC} < 0.6$, no discrimination [46]. Higher accuracy, sensitivity, specificity, and kappa are better.

4. Results

4.1. Demographics of the Study Population

Table 1 presents the patients' demographic characteristics and underlying diseases who underwent 3-vessel PCI for the first time. Patients who fulfilled the criteria from 1 January 2007 to 31 December 2009 were selected from the NHIRD. The mean follow-up periods were approximately 12 years in the survival group and approximately five years in the death group ($p < 0.001$). Men had a higher survival rate ($p < 0.001$).

Table 1. Baseline characteristics of patients who underwent 3-vessel PCI.

Variables			PCI (n = 2337)		p-Value
			Survive (n = 1188)	Death (n = 1149)	
Survival time (mean, SD) in years			11.34 (0.87)	4.57 (3.51)	<0.001
Baseline					
X1	Sex	Male Female	941 (79.21) 247 (20.79)	765 (66.58) 384 (33.42)	<0.001
X2	Age (mean, SD)		60.09 (10.18)	71.18 (10.24)	<0.001
X3	CCIS stratified	0	185 (15.57)	58 (5.05)	<0.001
		1	340 (28.62)	134 (11.66)	
		2	284 (23.91)	197 (17.15)	
		3	180 (15.15)	193 (16.8)	
		4	90 (7.58)	165 (14.36)	
		5	47 (3.96)	136 (11.84)	
		6+	62 (5.22)	266 (23.15)	
CCI score (mean, SD)			2.09 (1.80)	3.85 (2.50)	<0.001
One year before variables					
X4	Previous CABG surgery		7 (0.59)	13 (1.13)	0.1549
X5	Length of ICU stay (Days)		0.16 (1.69)	0.94 (6.83)	<0.001
X6	Blood transfusion (Bags)		0.12 (0.85)	0.88 (3.82)	<0.001
X7	Mechanical ventilation (Days)		0.12 (1.05)	0.87 (11.25)	0.0244
Current surgical variables					
X8	Blood transfusion (Bags)		0.17 (1.55)	3.42 (15.78)	<0.001
X9	Mechanical ventilation (Days)		0.13 (1.23)	3.45 (24.23)	<0.001
X10	Length of ICU stay (Days)		0.23 (2.07)	0.65 (4.80)	0.0063
X11	Length of Stay (Days)		4.29 (4.74)	9.85 (12.30)	<0.001
Underlying diseases					
X12	Hypertension		52 (4.38)	73 (6.35)	0.0338
X13	Hyperlipidemia		850 (71.55)	470 (40.91)	<0.001
X14	Hyperuricemia		52 (4.38)	30 (2.61)	0.0204
X15	Diabetes		499 (42)	644 (56.05)	<0.001
X16	Liver cirrhosis		6 (0.51)	27 (2.35)	<0.001
X17	COPD		108 (9.09)	228 (19.84)	<0.001
X18	Skin bone		128 (10.77)	158 (13.75)	0.0282
X19	Stroke		105 (8.84)	228 (19.84)	<0.001
X20	Gout		156 (13.13)	166 (14.45)	0.3561
X21	Biliary stone		27 (2.27)	25 (2.18)	0.8738
X22	HBV		39 (3.28)	21 (1.83)	0.0262
X23	HCV		19 (1.6)	24 (2.09)	0.3788
Heart-related characteristics					
X24	AF		32 (2.69)	105 (9.14)	<0.001
X25	CHF		175 (14.73)	537 (46.74)	<0.001
X26	PVD		49 (4.12)	102 (8.88)	<0.001
X27	ACS		487 (40.99)	654 (56.92)	<0.001
X28	Malignant dysrhythmia		21 (1.77)	80 (6.96)	<0.001
X29	Cardiogenic shock		11 (0.93)	87 (7.57)	<0.001
Renal-Related characteristics					
X30	KD		79 (6.65)	389 (33.86)	<0.001
X31	CRF		35 (2.95)	296 (25.76)	<0.001
X32	AKF		12 (1.01)	112 (9.75)	<0.001

Table 1. Cont.

Variables			PCI (n = 2337)		p-Value
			Survive (n = 1188)	Death (n = 1149)	
			N (%)		
Infection-related characteristics					
X33	Septicemia		412 (34.68)	439 (38.21)	0.0765
X34	Lower RTI		77 (6.48)	260 (22.63)	<0.001
X35	GTI		123 (10.35)	247 (21.5)	<0.001
Bleeding-related characteristics					
X36	Intracranial bleeding		13 (1.09)	28 (2.44)	0.0135
X37	TIA		108 (9.09)	233 (20.28)	<0.001
X38	Gastrointestinal bleeding		62 (5.22)	133 (11.58)	<0.001
X39	Major bleeding		44 (3.7)	52 (4.53)	0.3169
Hospital variables					
X40	Hospital status	Medical center	821 (69.11)	700 (60.92)	<0.001
		Non-medical center	367 (30.89)	449 (39.08)	
X41	Hospital ownership	Public hospital	361 (30.39)	332 (28.89)	0.4297
		Private hospital	827 (69.61)	817 (71.11)	
X42	Hospital area type	Central	197 (16.58)	234 (20.37)	<0.001
		Northern	647 (54.46)	467 (40.64)	
		Southern	332 (27.95)	407 (35.42)	
		Eastern	12 (1.01)	41 (3.57)	

ACS, acute coronary syndrome; AKF, acute kidney failure; AF, atrial fibrillation; CABG, coronary artery bypass grafting; CAD, coronary artery disease; CCI, Charlson Comorbidity Index; CHF, congestive heart failure; KD, kidney disease; COPD, chronic obstructive pulmonary disease; CRF, chronic renal failure; GTI, gastrointestinal infection; ICU, intensive care unit; LOS, length of stay; Lower RTI, lower respiratory tract infections; PCI, percutaneous coronary intervention; PVD, peripheral vascular disease; SD, standard deviation; TIA, transient ischemic attack.

Some patient characteristics were significantly different between the survival and death groups. The mean ages were 60.09 years and 71.18 years ($p < 0.001$), and the mean CCI scores were 2.09 ± 1.80 points and 3.85 ± 2.50 points ($p < 0.001$), respectively. In the year before PCI, the mean length of ICU stay (0.16 ± 1.69 vs. 0.94 ± 6.83 days, $p < 0.001$), blood transfusion (0.12 ± 0.85 vs. 0.88 ± 3.82 , $p < 0.001$), and days on mechanical ventilation (0.12 ± 1.05 vs. 0.87 ± 11.25 , $p = 0.0244$) were significantly different for the surviving and death groups. The surgical variables, including average bags of blood transfused (0.17 ± 1.55 vs. 3.42 ± 5.78 , $p < 0.001$), days on mechanical ventilation (0.13 ± 1.23 vs. 3.45 ± 24.23 , $p < 0.001$), length of ICU stay (0.23 ± 2.07 vs. 0.65 ± 4.80 , $p = 0.0063$), and length of hospital stay (4.29 ± 4.74 vs. 9.85 ± 12.30 days, $p < 0.001$) were all significantly different.

Most underlying diseases were significantly different for the survival and death groups, including hypertension (4.38% vs. 6.35%, $p = 0.0338$), hyperlipidemia (71.55% vs. 40.91%, $p < 0.001$), hyperuricemia (4.38% vs. 2.61%, $p = 0.0204$), diabetes (42% vs. 56.05%, $p < 0.001$), liver cirrhosis (0.51% vs. 2.35%, $p < 0.001$), COPD (9.09% vs. 19.84%, $p < 0.001$), skin and bone diseases (10.77% vs. 13.75%, $p = 0.0282$), stroke (8.84% vs. 19.84%, $p < 0.001$), and HBV (3.28% vs. 1.83%, $p = 0.0262$). The baseline heart-related characteristics were significantly different for the two groups, including AF (2.69% vs. 9.14%, $p < 0.001$), CHF (14.73% vs. 46.74%, $p < 0.001$), PVD (4.12% vs. 8.88%, $p < 0.001$), ACS (40.99% vs. 56.92%, $p < 0.001$), malignant dysrhythmia (1.77% vs. 6.96%, $p < 0.001$), and cardiogenic shock (0.93% vs. 7.57%, $p < 0.001$). The following renal-related characteristics were significantly

different between the survival and death groups: kidney disease (6.65% vs. 33.86%, $p < 0.001$), CRF (2.95% vs. 25.76%, $p < 0.001$), and AKF (1.01% vs. 9.75%, $p < 0.001$). Significantly different infection-related characteristics between the two groups included lower RTI (6.48% vs. 22.63%, $p < 0.001$) and GTI (10.35% vs. 21.5%, $p < 0.001$). Significantly different bleeding-related characteristics between the survival and death groups included intracranial bleeding (1.09% vs. 2.44%, $p = 0.0135$), TIA (9.09% vs. 20.28%, $p < 0.001$), and gastrointestinal bleeding (5.22% vs. 11.58%, $p < 0.001$). Hospital status and area were significantly different for the two groups.

4.2. Feature Importance

Different algorithms have different calculations. Table 2 presents the ranking of the selected variables by the six machine learning methods. LGR selected a total of 32 variables, 6 by CART, 10 by MARS, 26 by RF, and 13 by XGboost. A larger value means higher importance.

Table 2. Ranking of the selected variable by the six machine learning methods.

Methods		LGR (N = 32)	CART (N = 6)	MARS (N = 10)	RF (N = 26)	XGboost (N = 13)	Average
Baseline							
X1	Sex	4	-	-	15	-	3.8
X2	Age	32	31	32	32	32	31.8
X3	CCI scores	-	30	29	29	29	24.2
One year before variables							
X4	Previous CABG surgery	3	-	-	-	-	0.6
X5	Length of ICU stay (Days)	8	-	-	8	-	3.2
X6	Blood transfusion (Bags)	11	-	-	23	-	6.8
X7	Mechanical ventilation (Days)	-	-	-	16	-	3.2
Current surgical variables							
X8	Blood transfusion (Bags)	12	28	24	29	28	24.2
X9	Mechanical ventilation (Days)	16	-	-	25	24	13
Comorbidities							
X12	Hypertension	14	-	-	-	-	2.8
X13	Hyperlipidemia	28	-	27	26	26	21.4
X14	Hyperuricemia	23	-	-	-	-	4.6
X15	Diabetes	29	-	26	17	25	19.4
X17	COPD	20	-	-	21	-	8.2
X18	Skin bone	2	-	-	-	-	0.4
X19	Stroke	10	-	-	18	-	5.6
X20	Gout	-	-	-	7	-	1.4
X22	HBV	19	-	-	-	-	3.8
X23	HCV	18	-	-	-	-	3.6
Heart-related characteristics							
X24	AF	15	-	-	10	-	5
X25	CHF	31	29	30	30	30	30
X26	PVD	6	-	-	-	-	1.2
X27	ACS	22	-	25	22	23	18.4
X28	Malignant dysrhythmia	27	-	23	11	20	16.2
X29	Cardiogenic shock	25	-	28	12	22	17.6
Renal-related characteristics							
X30	KD	7	32	-	28	27	18.8
X31	CRF	30	27	31	27	29	28.8
X32	AKF	26	-	-	20	-	9.2

Table 2. Cont.

Methods		LGR (N = 32)	CART (N = 6)	MARS (N = 10)	RF (N = 26)	XGboost (N = 13)	Average
Infection-related characteristics							
X33	Septicemia	-	-	-	9	-	1.8
X34	Lower RTI	13	-	-	24	-	7.4
X35	GTI	-	-	-	14	-	2.8
Bleeding-related characteristics							
X36	Intracranial bleeding	17	-	-	-	-	3.4
X37	TIA	9	-	-	19	-	5.6
X38	Gastrointestinal bleeding	21	-	-	12	21	10.8
X39	Major bleeding	5	-	-	-	-	1

ACS, acute coronary syndrome; AKF, acute kidney failure; AF, atrial fibrillation; CAD, coronary artery disease; CABG, coronary artery bypass grafting; CART, classification and regression tree; CCI, Charlson Comorbidity Index; CHF, congestive heart failure; KD, kidney disease; COPD, chronic obstructive pulmonary disease; CRF, chronic renal failure; GTI, gastrointestinal infection; ICU, intensive care unit; LGR, logistic regression; LOS, length of stay; Lower RTI, lower respiratory tract infections; MARS, multivariate adaptive regression splines; PVD, peripheral vascular disease; RF, random forests; SD, standard deviation; TIA, transient ischemic attack. Xgboost, extreme gradient boosting.

Figure 2 presents the top 15 important risk factors ranked corresponding to the six machine learning models. The rankings of each selected variable in the six prediction models were as follows: age (which contributed the most to this model), CHF, CRF, CCI scores, hyperlipidemia, diabetes, KD, ACS, cardiogenic shock, malignant dysrhythmia, average days on a ventilator during recent surgery, gastrointestinal bleeding, AKF, and COPD. These 15 features can be considered the main variables for the prediction model. These rankings can give physicians an overview of the patient before or after the surgery. Furthermore, these variables were used to develop the ensemble learning prediction model.

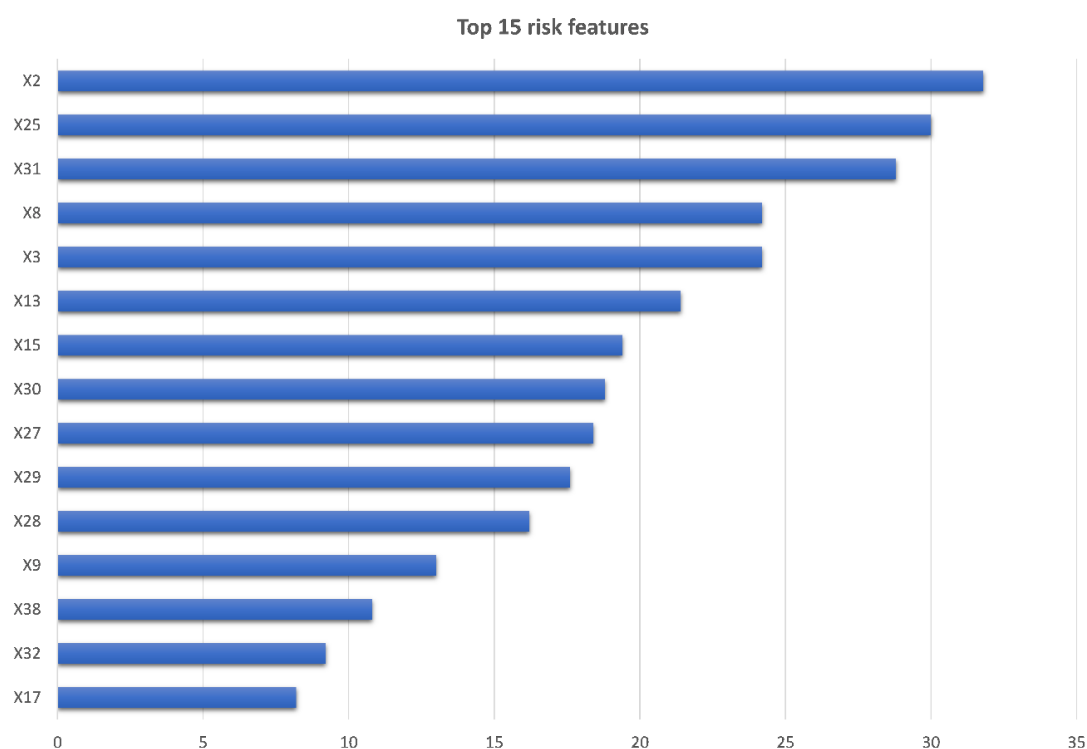


Figure 2. Results of the top 15 risk factors ranked in order.

Figure 3 shows the decision rules for predicting mortality among patients who underwent 3-vessel PCI for the first time, based on the three important features of the CART

model. Here, 86% of the patients aged <68 years and with a history of CRF died, whereas 78% of those without CRF died. Further, 94% of the patients aged 68–75 years with CRF died, and 76% of those without CRF but with CHF died. Finally, 75% of the patients aged >75 years died.

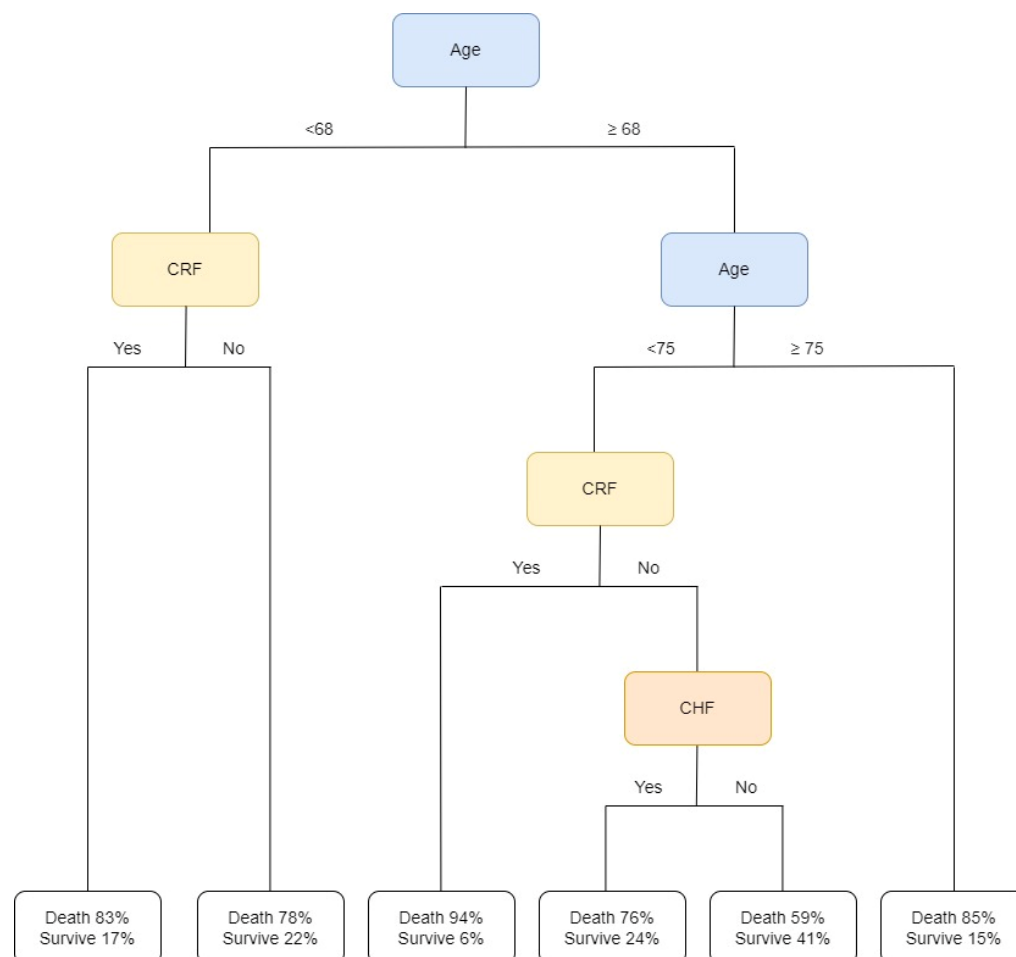


Figure 3. The survival prediction results in decision tree model analysis for PPCI patients. CHF, congestive heart failure; CRF, chronic renal failure.

4.3. Model Comparison

The ensemble learning method with majority voting was used to predict the survival of patients who underwent 3-vessel PCI for the first time. Overall, 15 risk factors were selected to develop single machine and ensemble learning models with an accuracy of 88.7% and an AUC of 90.0% in the ensemble learning model (Table 3).

Table 3. Accuracy of the single and ensemble learning prediction models after feature selection.

Methods	Single Machine Learning Models					Ensemble Learning Model
	LGR	CART	RF	MARS	XGboost	
Accuracy	0.752	0.778	0.810	0.761	0.739	0.887
Kappa	0.105	0.552	0.620	0.517	0.479	0.241
Sensitivity	0.833	0.684	0.844	0.640	0.764	0.833
Specificity	0.750	0.864	0.778	0.872	0.716	0.888
AUC	0.839	0.777	0.887	0.821	0.805	0.900
F1 Score	0.855	0.824	0.779	0.823	0.714	0.939

5. Discussion

Because of the tremendous advances in PCI instruments and technology, more patients with multivessel CAD prefer to undergo PCI instead of CABG. Before deciding whether or not to accept surgery, the doctor's clinical experience and SYNTAX score are used to make a preliminary judgment for the patient, and the SYNTAX is based on the severity of the blood vessels. SYNTAX score, which is a score according to coronary artery morphology, was used for the strategy of PCI or CABG for patients with triple vessel disease. For patients with low SYNTAX scores, PCI is an acceptable revascularization strategy, although at a price of significantly higher rates of repetitive revascularization [47]. The SYNTAX score and both are useful determinants for predicting hard clinical events (HCE: death, nonfatal myocardial infarction, and stroke). The incidence of HCE at three years significantly differed according to Clinical SYNTAX score (High, 20.2%; Intermediate, 1.2%; and Low, 6.0%; log-rank $p < 0.001$), but not according to SYNTAX score (High, 14.0%; Intermediate, 5.8%; and Low, 7.3%; log-rank $p = 0.13$) [48].

It is dangerous if we cannot judge only by looking at a single indicator for major decisions concerning life and death. Disease history and surgical factors are also useful for predicting patient prognoses after an intervention. Based on the risk features, we constructed an ensemble machine learning model using five well-known machine learning models and used majority voting to develop the best prediction model.

Renal insufficiency is a risk factor for poor long-term survival in patients who underwent coronary angiography [49], whereas age affects early and late survival independently after PCI [50]. Moreover, a history of CHF, acute myocardial infarction, cardiogenic shock, and renal disease is significantly associated with mortality [20,51,52]. Morbidity and mortality for PCI are high for patients with CAD and kidney disease [49,51,53–55]. The ISCHEMIA CKD trial included 777 patients with advanced renal insufficiency (eGFR < 30 mL/min) in the context of the larger ISCHEMIA trial population. An early routine invasive strategy failed to reduce the incidence of death or myocardial infarction, and an excess of stroke, death, or the initiation of dialysis was observed more than medical therapy alone [56]. CHF is a strong predictor of adverse outcomes after PCI [57]. Bleeding events were also associated with early and late mortality [58]. Some models were developed using elderly patient databases or fewer data to capture the important risk factors and prepare angiographic and mortality prediction models [18,52,59].

Moreover, feature selection can help physicians understand the causes of a disease; this method has been successfully applied in the medical industry [60]. The top 15 risk factors were identified using five single different machine learning models, and the rank of the scores for each variable was averaged. Figure 2 presents the ranking of important variables for predicting the outcome of PCI; age, CHF, and CRF are the top predictors of mortality, which is consistent with previous studies. Moreover, disease history and surgical risk factors were also identified as risk factors in this study. Although the safety of PCI has improved substantially, severely ill patients with different underlying diseases or complications who underwent 3-vessel PCI may experience prolonged hospital stays [58]. Figure 3 demonstrates the rules of the decision tree: patients aged 68–75 years with CRF will possibly have a 94% death rate. History of CRF should thus be confirmed before undergoing 3-vessel PCI for the first time; CABG or staged PTCA may be the preferred strategy for such patients.

This study demonstrated that feature selection, applying ensemble learning, and the presence of fewer risk factors could help achieve higher accuracy and identify factors affecting survival. This research provides a good basis and multiple-stage framework for developing better survival prediction models for patients undergoing 3-vessel PCI for the first time. An increasing number of patients are now opting for PCI as the treatment of choice for multivessel CAD. Physicians must consider the risk of multivessel PCI for each patient. The present study demonstrated that using feature selection and ensemble learning, fewer risk factors could achieve a higher accuracy rate and identify the factors affecting survival. This research provides useful suggestions and multiple-stage frameworks for

constructing a better survival prediction model for patients undergoing 3-vessel PCI for the first time, and physicians could make a more accurate choice.

6. Conclusions

Mortality prediction models are useful tools for physicians in clinical detection, and the prediction model's quality must be controlled and improved. This research provides a good basis and multiple-stage framework for developing better survival prediction models for patients undergoing 3-vessel PCI for the first time. In addition to identifying the 15 variables that affect survival, we used decision trees for inferences to predict patient mortality. This research compared different single and ensemble machine learning techniques. It can help physicians provide effective decisions and early healthcare management for patients older than 68 years with a history of kidney disease and improve the prediction model quality.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/app12168135/s1>, Table S1. Summary of the model parameter.

Author Contributions: Conceptualization, K.-Y.C. and M.C.; data curation, Y.-C.H.; formal analysis, Y.-C.H.; methodology, K.-Y.C. Y.-C.H.; project administration, K.-Y.C., C.-K.L., Y.-C.L. and M.C.; resources, M.C.; supervision, S.-J.L.; validation, K.-Y.C. and S.-J.L.; visualization, C.-K.L.; writing—original draft, Y.-C.H. and Y.-C.L.; writing—review & editing, K.-Y.C., Y.-C.H., C.-K.L., S.-J.L., Y.-C.L. and M.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Fu Jen Catholic University (A0109152) and the Ministry of Science and Technology (MOST109-2221-E030-011-MY3).

Institutional Review Board Statement: The study was conducted following the Declaration of Helsinki and approved by the Institutional Review Board of Fu Jen Catholic University (protocol code C108094 on 19 February 2020).

Informed Consent Statement: Patient consent was waived due to NHIRD providing a de-identification baseline and administrative information. The research project must review and process all NHIRD data before analysis.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sung, S.H.; Chen, T.C.; Cheng, H.M.; Lee, J.C.; Lang, H.C.; Chen, C.H. Comparison of Clinical Outcomes in Patients Undergoing Coronary Intervention with Drug-Eluting Stents or Bare-Metal Stents: A Nationwide Population Study. *Acta Cardiol. Sin.* **2017**, *33*, 10–19. [CrossRef] [PubMed]
2. Hsieh, M.H.; Lin, S.Y.; Lin, C.L.; Hsieh, M.J.; Hsu, W.H.; Ju, S.W.; Lin, C.C.; Hsu, C.Y.; Kao, C.H. A fitting machine learning prediction model for short-term mortality following percutaneous catheterization intervention: A nationwide population-based study. *Ann. Transl. Med.* **2019**, *7*, 732. [CrossRef] [PubMed]
3. Ellis, S.G.; Shishehbor, M.H.; Kapadia, S.R.; Lincoff, A.M.; Nair, R.; Whitlow, P.L.; Bajzer, C.T.; Cho, L.L.; Tuzcu, E.M.; Raymond, R.; et al. Enhanced prediction of mortality after percutaneous coronary intervention by consideration of general and neurological indicators. *JACC Cardiovasc. Interv.* **2011**, *4*, 442–448. [CrossRef] [PubMed]
4. Hanratty, C.G.; Koyama, Y.; Rasmussen, H.H.; Nelson, G.I.; Hansen, P.S.; Ward, M.R. Exaggeration of nonculprit stenosis severity during acute myocardial infarction: Implications for immediate multivessel revascularization. *J. Am. Coll. Cardiol.* **2002**, *40*, 911–916. [CrossRef]
5. Sasako, Y. Comparison of long-term prognosis between medical therapy PTCA and CABG for multiple coronary vessel disease. *Nihon Geka Gakkai Zasshi* **1996**, *97*, 215–219. [PubMed]
6. Sunagawa, G.; Komiya, T.; Tamura, N.; Sakaguchi, G.; Kobayashi, T.; Murashita, T. Coronary artery bypass surgery is superior to percutaneous coronary intervention with drug-eluting stents for patients with chronic renal failure on hemodialysis. *Ann. Thorac. Surg.* **2010**, *89*, 1896–1900; discussion 1900. [CrossRef]
7. Yap, C.H.; Yan, B.P.; Akowuah, E.; Dinh, D.T.; Smith, J.A.; Shardey, G.C.; Tatoulis, J.; Skillington, P.D.; Newcomb, A.; Mohajeri, M.; et al. Does prior percutaneous coronary intervention adversely affect early and mid-term survival after coronary artery surgery? *JACC Cardiovasc. Interv.* **2009**, *2*, 758–764. [CrossRef]
8. Malenka, D.J.; Leavitt, B.J.; Hearne, M.J.; Robb, J.F.; Baribeau, Y.R.; Ryan, T.J.; Helm, R.E.; Kellett, M.A.; Dauerman, H.L.; Dacey, L.J.; et al. Comparing long-term survival of patients with multivessel coronary disease after CABG or PCI: Analysis of BARI-like patients in northern New England. *Circulation* **2005**, *112*, I371–I376. [CrossRef]

9. Lin, W.C.; Chen, C.W.; Lu, C.L.; Lai, W.W.; Huang, M.H.; Tsai, L.M.; Li, C.Y.; Lai, C.H. The association between recent hospitalized COPD exacerbations and adverse outcomes after percutaneous coronary intervention: A nationwide cohort study. *Int. J. Chronic Obstr. Pulm. Dis.* **2019**, *14*, 169–179. [\[CrossRef\]](#)
10. Yin, W.H.; Lu, T.H.; Chen, K.C.; Cheng, C.F.; Lee, J.C.; Liang, F.W.; Huang, Y.T.; Yang, L.T. The temporal trends of incidence, treatment, and in-hospital mortality of acute myocardial infarction over 15 years in a Taiwanese population. *Int. J. Cardiol.* **2016**, *209*, 103–113. [\[CrossRef\]](#)
11. Halkin, A.; Grines, C.L.; Cox, D.A.; Garcia, E.; Mehran, R.; Tchong, J.E.; Griffin, J.J.; Guagliumi, G.; Brodie, B.; Turco, M.; et al. Impact of intravenous beta-blockade before primary angioplasty on survival in patients undergoing mechanical reperfusion therapy for acute myocardial infarction. *J. Am. Coll. Cardiol.* **2004**, *43*, 1780–1787. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Liu, E.S.; Hung, C.C.; Chiang, C.H.; Chang, C.H.; Cheng, C.C.; Kuo, F.Y.; Mar, G.Y.; Huang, W.C. Comparison of Different Timing of Multivessel Intervention During Index-Hospitalization for Patients With Acute Myocardial Infarction. *Front. Cardiovasc. Med.* **2021**, *8*, 639750. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Marenzi, G.; Assanelli, E.; Campodonico, J.; Lauri, G.; Marana, I.; De Metrio, M.; Moltrasio, M.; Grazi, M.; Rubino, M.; Veglia, F.; et al. Contrast volume during primary percutaneous coronary intervention and subsequent contrast-induced nephropathy and mortality. *Ann. Intern. Med.* **2009**, *150*, 170–177. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Chen, S.W.; Chang, C.H.; Lin, Y.S.; Wu, V.C.; Chen, D.Y.; Tsai, F.C.; Hung, M.J.; Chu, P.H.; Lin, P.J.; Chen, T.H. Effect of dialysis dependence and duration on post-coronary artery bypass grafting outcomes in patients with chronic kidney disease: A nationwide cohort study in Asia. *Int. J. Cardiol.* **2016**, *223*, 65–71. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Niles, N.W.; McGrath, P.D.; Malenka, D.; Quinton, H.; Wennberg, D.; Shubrooks, S.J.; Tryzelaar, J.F.; Clough, R.; Hearne, M.J.; Hernandez, F.; et al. Survival of patients with diabetes and multivessel coronary artery disease after surgical or percutaneous coronary revascularization: Results of a large regional prospective study. *J. Am. Coll. Cardiol.* **2001**, *37*, 1008–1015. [\[CrossRef\]](#)
16. Kuno, T.; Mikami, T.; Sahashi, Y.; Numasawa, Y.; Suzuki, M.; Noma, S.; Fukuda, K.; Kohsaka, S. Machine learning prediction model of acute kidney injury after percutaneous coronary intervention. *Sci. Rep.* **2022**, *12*, 749. [\[CrossRef\]](#)
17. Farkouh, M.E.; Domanski, M.; Dangas, G.D.; Godoy, L.C.; Mack, M.J.; Siami, F.S.; Hamza, T.H.; Shah, B.; Stefanini, G.G.; Sidhu, M.S.; et al. Long-Term Survival Following Multivessel Revascularization in Patients With Diabetes: The FREEDOM Follow-On Study. *J. Am. Coll. Cardiol.* **2019**, *73*, 629–638. [\[CrossRef\]](#)
18. Peterson, E.D.; Dai, D.; DeLong, E.R.; Brennan, J.M.; Singh, M.; Rao, S.V.; Shaw, R.E.; Roe, M.T.; Ho, K.K.; Klein, L.W.; et al. Contemporary mortality risk prediction for percutaneous coronary intervention: Results from 588,398 procedures in the National Cardiovascular Data Registry. *J. Am. Coll. Cardiol.* **2010**, *55*, 1923–1932. [\[CrossRef\]](#)
19. MacKenzie, T.A.; Malenka, D.J.; Olmstead, E.M.; Piper, W.D.; Langner, C.; Ross, C.S.; O'Connor, G.T.; Northern New England Cardiovascular Disease Study Group. Prediction of survival after coronary revascularization: Modeling short-term, mid-term, and long-term survival. *Ann. Thorac. Surg.* **2009**, *87*, 463–472. [\[CrossRef\]](#)
20. Hamburger, J.N.; Walsh, S.J.; Khurana, R.; Ding, L.; Gao, M.; Humphries, K.H.; Carere, R.; Fung, A.Y.; Mildenderger, R.R.; Simkus, G.J.; et al. Percutaneous coronary intervention and 30-day mortality: The British Columbia PCI risk score. *Catheter. Cardiovasc. Interv.* **2009**, *74*, 377–385. [\[CrossRef\]](#)
21. Duggal, B.; Subramanian, J.; Duggal, M.; Singh, P.; Rajivlochan, M.; Saunik, S.; Desiraju, K.; Avhad, A.; Ram, U.; Sen, S.; et al. Survival outcomes post percutaneous coronary intervention: Why the hype about stent type? Lessons from a healthcare system in India. *PLoS ONE* **2018**, *13*, e0196830. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Faxon, D.P.; Ghalilili, K.; Jacobs, A.K.; Ruocco, N.A.; Christellis, E.M.; Kellett, M.A., Jr.; Varrichione, T.R.; Ryan, T.J. The degree of revascularization and outcome after multivessel coronary angioplasty. *Am. Heart J.* **1992**, *123*, 854–859. [\[CrossRef\]](#)
23. Bourassa, M.; Yeh, W.; Holubkov, R.; Sopko, G.; Detre, K. Long-term outcome of patients with incomplete vs complete revascularization after multivessel PTCA: A report from the NHLBI PTCA Registry. *Eur. Heart J.* **1998**, *19*, 103–111. [\[CrossRef\]](#) [\[PubMed\]](#)
24. O'Connor, G.T.; Malenka, D.J.; Quinton, H.; Robb, J.F.; Kellett, M.A.; Shubrooks, S.; Bradley, W.A.; Hearne, M.J.; Watkins, M.W.; Wennberg, D.E. Multivariate prediction of in-hospital mortality after percutaneous coronary interventions in 1994–1996. *J. Am. Coll. Cardiol.* **1999**, *34*, 681–691. [\[CrossRef\]](#)
25. Chowdhary, S.; Ivanov, J.; Mackie, K.; Seidelin, P.H.; Džavík, V. The Toronto score for in-hospital mortality after percutaneous coronary interventions. *Am. Heart J.* **2009**, *157*, 156–163. [\[CrossRef\]](#)
26. Hizoh, I.; Domokos, D.; Banhegyi, G.; Becker, D.; Merkely, B.; Ruzsa, Z. Mortality prediction algorithms for patients undergoing primary percutaneous coronary intervention. *J. Thorac. Dis.* **2020**, *12*, 1706. [\[CrossRef\]](#)
27. Hizoh, I.; Gulyas, Z.; Domokos, D.; Banhegyi, G.; Majoros, Z.; Major, L.; Ratkai, T.; Kiss, R.G. A novel risk model including vascular access site for predicting 30-day mortality after primary PCI: The ALPHA score. *Cardiovasc. Revascularizat. Med.* **2017**, *18*, 33–39. [\[CrossRef\]](#)
28. de Mulder, M.; Gitt, A.; van Domburg, R.; Hochadel, M.; Seabra-Gomes, R.; Serruys, P.W.; Silber, S.; Weidinger, F.; Wijns, W.; Zeymer, U. EuroHeart score for the evaluation of in-hospital mortality in patients undergoing percutaneous coronary intervention. *Eur. Heart J.* **2011**, *32*, 1398–1408. [\[CrossRef\]](#)
29. Huang, Y.C.; Li, S.J.; Chen, M.; Lee, T.S. The Prediction Model of Medical Expenditure Applying Machine Learning Algorithm in CABG Patients. *Healthcare* **2021**, *9*, 710. [\[CrossRef\]](#)

30. Ghosh, P.; Azam, S.; Jonkman, M.; Karim, A.; Shamrat, F.M.J.M.; Ignatious, E.; Shultana, S.; Beeravolu, A.R.; De Boer, F. Efficient Prediction of Cardiovascular Disease Using Machine Learning Algorithms With Relief and LASSO Feature Selection Techniques. *IEEE Access* **2021**, *9*, 19304–19326. [\[CrossRef\]](#)
31. Muthukrishnan, R.; Rohini, R. LASSO: A feature selection technique in predictive modeling for machine learning. In Proceedings of the 2016 IEEE International Conference on Advances in Computer Applications (ICACA), Coimbatore, India, 24 October 2016; pp. 18–20.
32. Liu, Z.; Chen, H.; Sun, X.; Chen, H. Data-Driven Real-Time Online Taxi-Hailing Demand Forecasting Based on Machine Learning Method. *Appl. Sci.* **2020**, *10*, 6681. [\[CrossRef\]](#)
33. Asman, S.H.; Ab Aziz, N.F.; Ungku Amirulddin, U.A.; Ab Kadir, M.Z.A. Decision Tree Method for Fault Causes Classification Based on RMS-DWT Analysis in 275 kV Transmission Lines Network. *Appl. Sci.* **2021**, *11*, 4031. [\[CrossRef\]](#)
34. Pourhomayoun, M.; Shakibi, M. Predicting mortality risk in patients with COVID-19 using machine learning to help medical decision-making. *Smart Health* **2021**, *20*, 100178. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Kang, I.-A.; Ngnamsie Njimboum, S.; Lee, K.-O.; Kim, J.-D. DCP: Prediction of Dental Caries Using Machine Learning in Personalized Medicine. *Appl. Sci.* **2022**, *12*, 3043. [\[CrossRef\]](#)
36. Park, M.; Jung, D.; Lee, S.; Park, S. Heatwave Damage Prediction Using Random Forest Model in Korea. *Appl. Sci.* **2020**, *10*, 8237. [\[CrossRef\]](#)
37. Khoshgoftaar, T.M.; Golawala, M.; Van Hulse, J. An empirical study of learning from imbalanced data using random forest. In Proceedings of the 19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2007), Patras, Greece, 29–31 October 2007; pp. 310–317.
38. Friedman, J.H. Multivariate adaptive regression splines. *Ann. Stat.* **1991**, *19*, 1–67. [\[CrossRef\]](#)
39. Sekhar Roy, S.; Roy, R.; Balas, V.E. Estimating heating load in buildings using multivariate adaptive regression splines, extreme learning machine, a hybrid model of MARS and ELM. *Renew. Sustain. Energy Rev.* **2018**, *82*, 4256–4268. [\[CrossRef\]](#)
40. Zhang, W.G.; Goh, A.T.C. Multivariate adaptive regression splines for analysis of geotechnical engineering systems. *Comput. Geotech.* **2013**, *48*, 82–95. [\[CrossRef\]](#)
41. Raza, K. Improving the prediction accuracy of heart disease with ensemble learning and majority voting rule. In *U-Healthcare Monitoring Systems*; Academic Press: Cambridge, MA, USA, 2019; pp. 179–196. [\[CrossRef\]](#)
42. Kim, S.Y.; Upneja, A. Majority voting ensemble with a decision trees for business failure prediction during economic downturns. *J. Innov. Knowl.* **2021**, *6*, 112–123. [\[CrossRef\]](#)
43. Matloob, F.; Ghazal, T.M.; Taleb, N.; Aftab, S.; Ahmad, M.; Khan, M.A.; Abbas, S.; Soomro, T.R. Software defect prediction using ensemble learning: A systematic literature review. *IEEE Access* **2021**, *9*, 98754–98771. [\[CrossRef\]](#)
44. Rokach, L. Ensemble-based classifiers. *Artif. Intell. Rev.* **2010**, *33*, 1–39. [\[CrossRef\]](#)
45. Rengaraju, U. Ensemble Learning Techniques—VotingClassifier. Available online: <https://medium.com/wids-mysore/ensemble-learning-techniques-votingclassifier-c4b38ee62129> (accessed on 29 June 2022).
46. Hu, Y.-H.; Chen, K.; Chang, I.-C.; Shen, C.-C. Critical predictors for the early detection of conversion from unipolar major depressive disorder to bipolar disorder: Nationwide population-based retrospective cohort study. *JMIR Med. Inform.* **2020**, *8*, e14278. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Head, S.J.; Davierwala, P.M.; Serruys, P.W.; Redwood, S.R.; Colombo, A.; Mack, M.J.; Morice, M.-C.; Holmes, D.R., Jr.; Feldman, T.E.; Stähle, E. Coronary artery bypass grafting vs. percutaneous coronary intervention for patients with three-vessel disease: Final five-year follow-up of the SYNTAX trial. *Eur. Heart J.* **2014**, *35*, 2821–2830. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Kashiwama, T.; Otsuji, S.; Takiuchi, S.; Asano, K.; Ibuki, M.; Hasegawa, K.; Ishibuchi, K.; Fujino, A.; Ishii, R.; Higashino, Y. A multidirectional approach to risk assessment in patients with three-vessel coronary artery disease undergoing percutaneous intervention. *J. Cardiol.* **2017**, *69*, 640–647. [\[CrossRef\]](#)
49. Hemmelgarn, B.R.; Ghali, W.A.; Quan, H.; Brant, R.; Norris, C.M.; Taub, K.J.; Knudtson, M.L. Poor long-term survival after coronary angiography in patients with renal insufficiency. *Am. J. Kidney Dis.* **2001**, *37*, 64–72. [\[CrossRef\]](#) [\[PubMed\]](#)
50. Taddei, C.F.G.; Weintraub, W.S.; Douglas, J.S., Jr.; Ghazzal, Z.; Mahoney, E.; Thompson, T.; King, S., III. Influence of age on outcome after percutaneous transluminal coronary angioplasty. *Am. J. Cardiol.* **1999**, *84*, 245–251. [\[CrossRef\]](#)
51. Rubenstein, M.H.; Harrell, L.C.; Sheynberg, B.V.; Schunkert, H.; Bazari, H.; Palacios, I.F. Are patients with renal failure good candidates for percutaneous coronary revascularization in the new device era? *Circulation* **2000**, *102*, 2966–2972. [\[CrossRef\]](#)
52. Klein, L.W.; Shaw, R.E.; Krone, R.J.; Brindis, R.G.; Anderson, H.V.; Block, P.C.; McKay, C.R.; Hewitt, K.; Weintraub, W.S.; American College of Cardiology National Cardiovascular Data Registry. Mortality after emergent percutaneous coronary intervention in cardiogenic shock secondary to acute myocardial infarction and usefulness of a mortality prediction model. *Am. J. Cardiol.* **2005**, *96*, 35–41. [\[CrossRef\]](#)
53. McCullough, P.A.; Soman, S.S.; Shah, S.S.; Smith, S.T.; Marks, K.R.; Yee, J.; Borzak, S. Risks associated with renal dysfunction in patients in the coronary care unit. *J. Am. Coll. Cardiol.* **2000**, *36*, 679–684. [\[CrossRef\]](#)
54. Ting, H.H.; Tahirkheli, N.K.; Berger, P.B.; McCarthy, J.T.; Timimi, F.K.; Mathew, V.; Rihal, C.S.; Hasdai, D.; Holmes, D.R., Jr. Evaluation of long-term survival after successful percutaneous coronary intervention among patients with chronic renal failure. *Am. J. Cardiol.* **2001**, *87*, 630–633. [\[CrossRef\]](#)
55. Szczech, L.A.; Reddan, D.N.; Owen, W.F., Jr.; Califf, R.; Racz, M.; Jones, R.H.; Hannan, E.L. Differential survival after coronary revascularization procedures among patients with renal insufficiency. *Kidney Int.* **2001**, *60*, 292–299. [\[CrossRef\]](#) [\[PubMed\]](#)

-
56. Caracciolo, A.; Scalise, R.F.M.; Ceresa, F.; Bagnato, G.; Versace, A.G.; Licordari, R.; Perfetti, S.; Lofrumento, F.; Irrera, N.; Santoro, D.; et al. Optimizing the Outcomes of Percutaneous Coronary Intervention in Patients with Chronic Kidney Disease. *J. Clin. Med.* **2022**, *11*, 2380. [[CrossRef](#)] [[PubMed](#)]
 57. Holper, E.M.; Blair, J.; Selzer, F.; Detre, K.M.; Jacobs, A.K.; Williams, D.O.; Vlachos, H.; Wilensky, R.L.; Coady, P.; Faxon, D.P. The impact of ejection fraction on outcomes after percutaneous coronary intervention in patients with congestive heart failure: An analysis of the National Heart, Lung, and Blood Institute Percutaneous Transluminal Coronary Angioplasty Registry and Dynamic Registry. *Am. Heart J.* **2006**, *151*, 69–75.
 58. Chhatrwalla, A.K.; Amin, A.P.; Kennedy, K.F.; House, J.A.; Cohen, D.J.; Rao, S.V.; Messenger, J.C.; Marso, S.P.; National Cardiovascular Data Registry. Association between Bleeding Events and In-hospital Mortality after Percutaneous Coronary Intervention. *JAMA* **2013**, *309*, 1022–1029. [[CrossRef](#)] [[PubMed](#)]
 59. Negassa, A.; Monrad, E.S.; Bang, J.Y.; Srinivas, V.S. Tree-structured risk stratification of in-hospital mortality after percutaneous coronary intervention for acute myocardial infarction: A report from the New York State percutaneous coronary intervention database. *Am. Heart J.* **2007**, *154*, 322–329. [[CrossRef](#)] [[PubMed](#)]
 60. Remeseiro, B.; Bolon-Canedo, V. A review of feature selection methods in medical applications. *Comput. Biol. Med.* **2019**, *112*, 103375. [[CrossRef](#)] [[PubMed](#)]