Journal of Universal Computer Science, vol. 32, no. 8 (2026), 1203-1222
submitted: 11/8/2025, accepted: 29/12/2025, appeared: 28/8/2026 CC BY 4.0

Software Development Model for Predicting Survival of Heart Failure Patient Using Multi-Algorithm Method

Zulkifli Zulkifli
(Informatics Engineering, Faculty of Technology and Informatics, Aisyah University, Indonesia
https://orcid.org/0000-0002-8702-0885, zulkifli@aisyahuniversity.ac.id)

Fitriana Fitriana
(Midwifery, Faculty of Health, Aisyah University, Indonesia
https://orcid.org/0000-0003-3230-2591, fitriana@aisyahuniversity.ac.id)

Panji Bintoro
(Software Engineering, Faculty of Technology and Informatics, Aisyah University, Indonesia
https://orcid.org/0000-0002-2955-1872, panjibintoro09@aisyahuniversity.ac.id)

Abstract: Heart failure (HF) is a major global health issue and a leading cause of hospitalization as well as mortality, specifically among the elderly. Accurate prediction and continuous monitoring of HF patient survival are essential for effective clinical decision-making. Therefore, this study aimed to develop a web-based software system that incorporates multi-algorithm, Naive Bayes (NB), Random Forest (RF), K-Nearest Neighbor (KNN), Support Vector Machine (SVM), and Neural Network (NN), to predict patient survival using 12 clinical variables. The software enabled efficient diagnosis and monitoring by generating predictive outcomes based on patient data. In the evaluation phase, the NB algorithm achieved the highest accuracy of 88.21%, followed by RF (85.11%), NN (82.36%), SVM (76.11%), and KNN (62%). The major originality of this study resided in combining multi-algorithm into a unified diagnostic platform, offering practical and low-cost decision support for healthcare providers. In contrast to previous studies that focused on developing standalone model, this analysis produced a functional application directly implemented in clinical environments. Relating to the analysis, future studies should aim to expand the dataset, incorporate real-time hospital data, and use explainable AI to improve interpretability, scalability, as well as trustworthiness in medical decision support systems.

Keywords: Heart Failure (HF), Multi-Algorithm, Prediction, Accuracy Level
Categories: D.2.3, I.1.2, G.2.3
DOI: 10.3897/jucs.168348

1 Introduction

Heart failure (HF) is a prevalent health condition affecting millions worldwide and remains the leading cause of hospitalization among elderly patient [Abdul-Samad et al., 2024]. Rather than being classified as a single disease, HF is better understood as a complex clinical syndrome. It impairs the ability of the heart to pump sufficient blood throughout the body, typically occurring when the ventricles cannot adequately fill with or eject blood [Mandal et al., 2021]. Common symptoms include shortness of breath, ankle swelling, and fatigue, while clinical signs may comprise elevated jugular venous pressure, crackles in the lungs, as well as peripheral edema [Ghaffar Nia et al., 2023].

In most cases, these manifestations come from structural or functional abnormalities of the heart or related systems.

HF is a serious condition with high morbidity and mortality rates. According to the European Society of Cardiology (ESC), approximately 26 million adults worldwide are living with the condition, with 3.6 million new cases diagnosed each year. Mortality rate is high, with 17–45% of patient dying in the first year after diagnosis, and most of the remainder deceasing in five years. The economic burden is also considerable, as managing HF accounts for about 1–2% of total healthcare expenditures, the majority of which are driven by recurrent hospitalizations [Tripoliti et al., 2017].

In 2019, the World Health Organization (WHO) reported that approximately 17.9 million deaths were caused by cardiovascular diseases (CVDs) or HF, totaling 32% of all global deaths. Major behavioral risk factors for heart disease include unhealthy eating habits, physical inactivity, tobacco use, and excessive alcohol consumption. Among environmental risk factors, air pollution plays a significant role. In the context of this study, the behavioral risk factors may lead to intermediate conditions such as elevated blood pressure, high blood glucose levels, increased blood lipid levels, as well as overweight and obesity [Huang et al., 2023]. These intermediate risk factors, which can be assessed in primary healthcare settings, are connected to a higher possibility of heart attacks, strokes, HF, and other related complications [Kerr & Lagen, 2021].

Diagnosing heart disease presents a challenge at the early stages for healthcare professionals, as the nature of these symptoms can often be misinterpreted as aging-related issues. Building on this topic, angiography is widely used by medical professionals for accurate diagnosis of coronary artery disease. High cost and limited availability of the method make the process difficult for low-income individuals to access [Petry, 2020]. The growing volume of medical data provides an opportunity to improve disease diagnosis. Moreover, advances in computer technology assist professionals in decision-making processes. Machine learning (ML) is an effective technology for transforming large datasets into valuable knowledge [Umer et al., 2022].

Developing software that can predict the survival of HF patient using a neural network algorithm method is essential to improve the monitoring system. The diagnostic and monitoring outcomes from this software can serve as an initial reference for diagnosing patient with lower costs and greater accuracy [AlDera & Othman, 2022].

Several studies by Ke Wang in 2021 combined ML with Shapley Additive exPlanations (SHAP) to explain individual risk prediction and provide an understanding to doctors about the causes of HF-related mortality after calculating all-cause mortality risks over three years [K. Wang et al., 2021]. In the development of software for monitoring HF patient, a literature review of studies related to multi-algorithm development, data mining, and ML is necessary to be conducted. A similar study was performed by Feda Anisah Makkiyah in 2024, which explained the development of applications and the implementation of a multi-algorithm for predicting diabetes status [Zulkifli et al., 2024]. Another relevant analysis by Muhammad Umer in 2022 focused on kidney failure patient care with an Internet of Things (IoT) method [Umer et al., 2022].

2 Related Works

The study by Chicco & Jurman [Chicco & Jurman, 2020] showed that ML could effectively predict the survival of HF patient using only two major indicators, namely serum creatinine and ejection fraction. This method enabled physicians focus personal prognosis on two crucial features. Using data from 299 HF patient in Faisalabad (Pakistan) collected between April and December 2015, the study confirmed the capability of ML to classify cardiovascular patient health records. However, the study was limited as it did not include physical information and occupational history of patient, which could potentially show additional risk factors. Further studies evaluated ML model to predict three-year mortality in HF patient caused by coronary heart disease. The XGBoost model achieved the highest accuracy and provided individualized risk assessments that were interpretable through SHAP values, supporting clinical decision-making. Among 1,562 patient studied, 210 (13.44%) died in three years. Following the discussion, XGBoost showed a hazard ratio of 10.351, which remained significant (5.343) after multivariable analysis. The study was limited as it only applied commonly used ML methods without comparison to validated risk calculators and prioritized the need for ML model capable of explaining the reasoning behind the prediction [K. Wang et al., 2021]. Furthermore, Saboor et al. [Saboor et al., 2022] investigated the use of ML algorithm for detecting and predicting heart disease, showing the efficiency compared to manual methods. The study evaluated various classifiers on a heart disease dataset and found that data standardization and hyperparameter tuning significantly improved accuracy by up to 8.78%. These results showed the importance of algorithm optimization in improving model performance. Early detection of heart disease is considered crucial for preventing HF and improving patient survival, signifying the need for reliable diagnostic methods. This recent study compared ML with conventional statistical model (CSM) in predicting 30-day readmission among patient with HF. The results showed that although both models had similar levels of discrimination, the Fine-Gray signified better calibration, implying the importance of accurately reporting calibration metrics in ML model. Using independent data from the EFFECT-1 study (1999–2001) for external validation, the Fine-Gray model achieved a c-statistic of 0.621, while RSF-CR accomplished 0.620. The limitation of the study was in the data period (2004–2005), which excluded newer medications and biomarkers, potentially affecting the accuracy. In general, accurate predictive model have the potential to improve patient outcomes and reduce healthcare costs, making the systems essential for clinical practice [Abdul-Samad et al., 2024].

The study by Xie et al. [Xie et al., 2025] explained the challenges in diagnosing HF and the potential of artificial intelligence (AI) to improve early detection as well as clinical decision-making. The study explored AI algorithm capable of transforming patient care through more in-depth data analysis. AI-based ECG analysis showed outstanding performance, achieving a C-statistic of 0.92 in detecting HF with reduced ejection fraction (HFrEF). However, the main limitation of AI was in its “black box” nature, which restricted interpretability and led to errors in causal inference. The study proposed a novel method that combined Angle Transformation (AT) and Long Short-Term Memory (LSTM) networks for the automatic classification of ECG signals, achieving high accuracy in detecting congestive heart failure (CHF) and arrhythmia (ARR). Using ECG data from the MIT-BIH and BIDMC databases, the AT + LSTM method achieved a success rate of 97.75%–98.97%. Consequently, previous methods

were considered less robust and accurate for real-time monitoring. The process showed the challenges in diagnosing HF with preserved ejection fraction (HFpEF), which affected many people worldwide. It prioritized the importance of early detection and treatment to prevent disease progression, as well as reduced the burden on both patient and healthcare systems. Statistical analysis using a 95% confidence interval from bootstrap sampling showed the robustness of the results. However, the categorical output with an "intermediate" classification complicated model comparisons and affected clinical applicability. Understanding these results was crucial for improving patient management and the effectiveness of HFpEF treatment [Akerman et al., 2025]. Following the discussion, the analysis evaluated various ML algorithm, including Random Forest (RF) and Decision Tree, to predict HF using datasets from Kaggle. After data preprocessing, the RF algorithm achieved an accuracy of 90.76%, showing its effectiveness in medical diagnosis. Based on the HF Prediction Dataset and the bi-state heart disease data from Kaggle, the algorithm recorded an accuracy of 93.1% for HF patient and 87.8% for healthy individuals. Considering that CVDs caused 17.9 million deaths globally in 2019, this study showed the importance of accurate predictive model for the early detection and prevention of HF [Coşkun, 2022].

Existing studies have primarily focused on developing an ML-based model for predicting HF patient survival. However, the model have not translated into the practical software that hospitals or doctors could use as an initial reference tool for diagnosis and prognosis [Giri et al., 2021]. To bridge the gap, the proposed software will be developed as a mobile web-based platform, incorporating 13 input variables, namely patient code, age, anemia, creatinine phosphokinase, diabetes, ejection fraction, high blood pressure, platelets, serum creatinine, serum sodium, sex, smoking, and time. The system will then generate a single output variable, signifying how a death event has occurred.

3 The Purpose Method

3.1 Study Framework

The section included stages of software development for the monitoring system and survival prediction of HF patient using a neural network algorithm. The processes were later evaluated for accuracy using Naive Bayes (NB), RF, KNN, SVM, and NN.

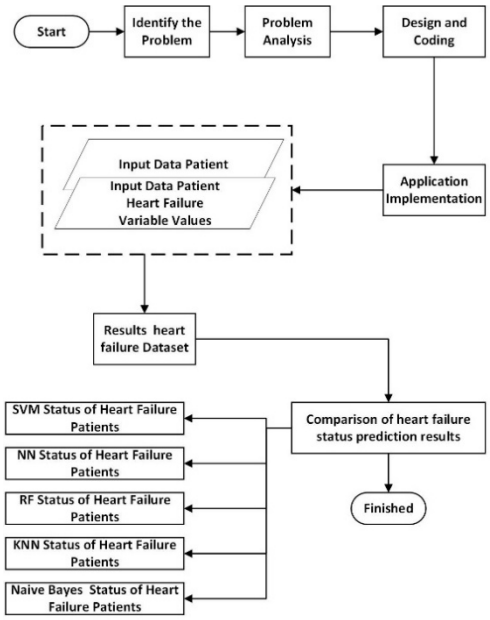

Figure 1: The stages of software development for a monitoring and survival prediction system for HF patient, which was later evaluated for accuracy using NB, RF, KNN, SVM, and NN algorithm.

Figure 1 showed the stages of software development for a monitoring and survival prediction system for HF patient. The first steps included identifying the issue, problem analysis, design, and coding. During this process, the application implementation included two processes, namely inputting patient data and HF variable values. The analysis produced a dataset of HF conditions and a comparison of its status prediction outcome using various ML model. The model comprised NB [Varshni et al., 2019], RF [Singh & Verma, 2018], KNN [Mallikharjuna Rao et al., 2023], SVM [Ismael & Şengür, 2021], and NN [Goyal, 2022] Status of HF Patient.

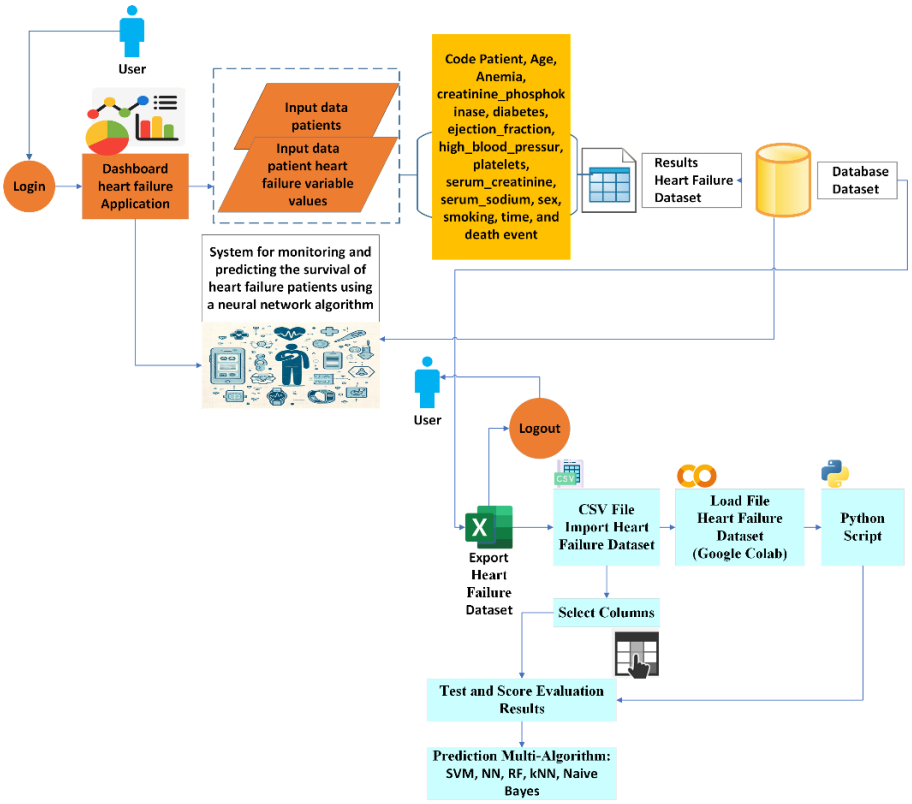

Figure 2: The study framework for the development of software for monitoring and predicting the survival of HF patient.

Figure 2 showed the study framework for the development of software for monitoring and predicting the survival of HF patient. The framework offered a high level of flexibility, as it was developed to meet specific needs, consisting of several stages, starting with user login. After successful authentication during the process, the application dashboard was shown. The user then input patient data and HF variable values into the model. The variables that were entered included patient code, age, anemia, creatinine phosphokinase, diabetes, ejection fraction, high blood pressure, platelets, serum creatinine, serum sodium, sex, smoking, and time. Based on the 12 input variables, the system generated a death event status from data processing using a neural network algorithm. In this process, the output data was generated formed HF dataset, which was stored in the database. The dataset in the database was exported as a CSV file, which was loaded into Google Colab for scripting using Python. To measure accuracy, the study used the orange widget tool during the analysis. From this process, multi-algorithm prediction was obtained, including NN [Cross et al., 1995], SVM [Aydin et al., 2007], KNN [Goyal, 2022a], RF [Dritsas & Trigka, 2022], and NB [Ali et al., 2023].

4 Result

Software development stages for HF patient monitoring and survival prediction system included generating a dataset from the patient input variable, which was used to predict patient survival. This process was conducted using NB, RF, KNN, SVM, and Neural Network algorithm, which were as follows.

During the analysis, problem identification was conducted at this stage. The identified problem was needed for the development of software for a monitoring and survival prediction system for HF patient to measure the accuracy of survival prediction. This development included 12 input data variables for the patient, including patient code, age, anemia, creatinine phosphokinase, diabetes, ejection fraction, high blood pressure, platelets, serum creatinine, serum sodium, sex, smoking, and time.

4.1 Problem Identification

In this stage, problem identification was performed, including:

a. The need for software that enabled a monitoring and survival prediction system for HF patient, supporting decision-making and policy implementation concerning patient care.
b. The developed model should be applicable in healthcare services.
c. Measuring the accuracy of survival prediction for HF patient using NB, RF, KNN, SVM, and Neural Network algorithm.

4.2 Problem Analysis

Engaging with tool users to ascertain software expectations from both physicians and healthcare institutions.

4.3 Result of Disease Detection

Software for the HF Patient Survival Monitoring System could be accessed at the following link: https://heartfailure.coi-uap.com/index.php. Figure 3 showed some screenshots of the HF Patient Survival Prediction Monitoring System.

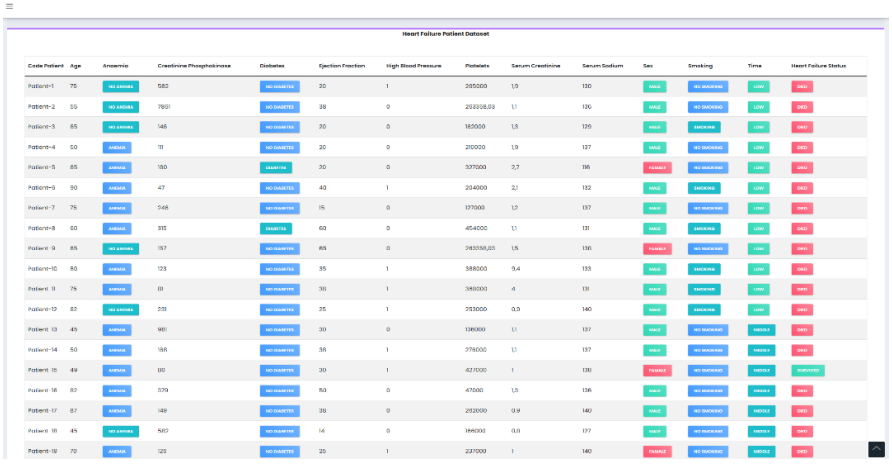

Figure 3: View of the monitoring system and prediction of survival of HF patient

Figure 3 showed the interface of the monitoring and survival prediction system for HF patient. This system had three access levels, namely doctor, hospital, and administrator.

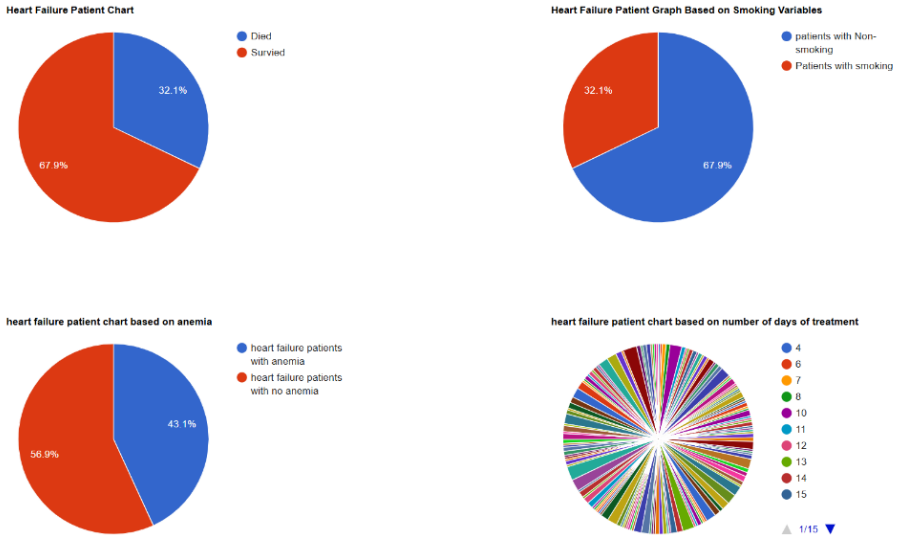

Figure 4: Testing time results with seven model

The interface was used by doctors and hospital staff to view as well as monitor the data of HF patient. The data included graphs based on anemia, smoking, and the duration of hospitalization.

4.4 Dataset Database

Table 1 showed the dataset during the process of measuring the accuracy of survival prediction for HF patient using NB, RF, KNN, SVM, and NN algorithm [Esteban et al., 2019].

1	2	3	4	5	6	7	8	9	10	11	12	13	14
Patient-01	75	0	582	0	20	1	265.000	1,9	130	1	4	1	Died
Patient-02	55	0	7.861	0	38	0	263.358	1,1	136	1	6	1	Died
Patient-03	65	0	146	0	20	0	162.000	1,3	129	1	7	1	Died
Patient-04	50	1	111	0	20	0	210.000	1,9	137	1	7	1	Died
Patient-05	65	1	160	1	20	0	327.000	2,7	116	0	8	1	Died
Patient-06	90	1	47	0	40	1	204.000	2,1	132	1	8	1	Died
Patient-07	75	1	246	0	15	0	127.000	1,2	137	1	10	1	Died
Patient-15	49	1	80	0	30	1	427.000	1	138	0	12	0	Survived
Patient-21	65	1	52	0	25	1	276.000	1,3	137	0	16	0	Survived
Patient-31	94	0	582	1	38	1	263.358	1,83	134	1	27	0	Survived

Note: 1=Code Patient, 2=Age, 3=Anaemia, 4=Creatinine Phosphokinase, 5=Diabetes, 6=Ejection Fraction, 7=High Blood Pressure, 8=Platelets, 9=Serum Creatinine, 10=Serum Sodium, 11=Sex, Smoking, 12=Time, 13=Death Event, and HF Status

Table 1: Performance metrics for each test class

4.5 Dataset Normalization with NN

Table 2 showed a sample of the dataset after normalization during the process. The transformation formula scaled the original test data values to a range between 0.1 and 0.9. This step was essential because the model used a sigmoid activation function, which produced outputs greater than 0 [Byun et al., 2019].

1	2	3	4	5	6	7	8	9	10	11	12	13	ANN		10
													a	b	
Patient-01	75	0	582	0	20	1	265.000	1,9	130	1	4	1	0.992	0.007	Died
Patient-02	55	0	7.86	0	38	0	263.358	1,1	136	1	6	1	0.245	0.754	Died
Patient-03	65	0	146	0	20	0	162.000	1,3	129	1	7	1	0.999	0.001	Died
Patient-04	50	1	111	0	20	0	210.000	1,9	137	1	7	1	0.999	0.001	Died
Patient-05	65	1	160	1	20	0	327.000	2,7	116	0	8	1	0.990	0.009	Died
Patient-06	90	1	47	0	40	1	204.000	2,1	132	1	8	1	0.999	0.001	Died
Patient-07	75	1	246	0	15	0	127.000	1,2	137	1	10	1	0.981	0.018	Died
Patient-15	49	1	80	0	30	1	427.000	1	138	0	12	0	0.999	0.001	Survived
Patient-21	65	1	52	0	25	1	276.000	1,3	137	0	16	0	0.999	0.001	Survived
Patient-31	94	0	582	1	38	1	263.358	1,83	134	1	27	0	0.995	0.001	Survived

Note: 1=Code Patient, 2=Age, 3=Anaemia, 4=Creatinine Phosphokinase, 5=Diabetes, 6=Ejection Fraction, 7=High Blood Pressure, 8=Platelets, 9=Serum Creatinine, 10=Serum Sodium, 11=Sex, Smoking, 12=Time, 13=Death Event, and HF Status

Table 2: Dataset normalization

4.6 Model Training and Methodology

The Adam optimization algorithm was used to train each layer of the NN model, with parameter adjustments applied during the process. The training configuration included 20 epochs, a validation split of 0.4, and a batch size of 16, applied across the entire dataset [H. Wang et al., 2021].

Epoch	Learning Rate	Hidden Layers	Inputs Layers	Output Layers	Error per epoch
500	0.2	4,6,3,4,5	12	1	0.2209665
500	0.3	4,6,3,4,5	12	1	0.2214913
1000	0.2	4,6,3,4,5	12	1	0.2180693
1000	0.3	4,6,3,4,5	12	1	0.2206641

Table 3: Performance of NN

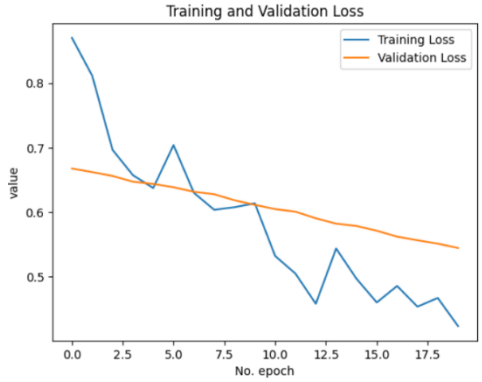

Figure 5: Performance of loss training and validation NN algorithm

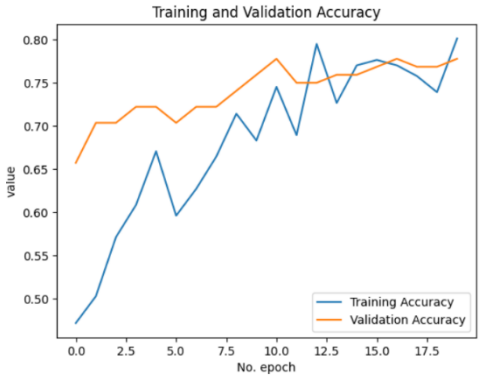

Figure 6: Performance of the accuracy training and validation NN algorithm.

The training and testing processes produced a training loss of 0.5461 and a prediction accuracy of 0.80, as shown in Figures 5 and 6, signifying the performance of the NN model. When tested using the Stochastic Gradient Descent (SGD) optimization method [Xu et al., 2021], with a batch size of 16, 20 epochs, and a validation split of 0.4, the model achieved a training loss of 0.4793 as well as an accuracy of 0.80, as presented in Figures 2 and 3. Furthermore, Table 4 showed the testing results for the SVM algorithm [Niu, 2018].

Model	Parameter Settings	Accuracy
SVM	(C=2; linear; scale; degree=3)	0.86
	(C=1; rbf; scale; degree=4)	0.90
	(C=1; linear; scale; degree=3)	0.86

Table 4: Performance of SVM

The results of the SVM method, using an RBF kernel [Meeradevi et al., 2025] with parameters C=1, degree = 4, and gamma = "scale," achieved an accuracy of 0.90, as shown in the previous table. Similarly, Table 5 showed the testing outcomes for the RF algorithm.

Model	Parameter Settings	Accuracy
RF	(n_estimators=200, random_state=3, max_depth=5, criterion='entropy', class_weight = None, verbose = 0)	0.93

Table 5: Performance of RF

The results of the RF method, configured with parameters max_depth = 5, random_state = 3, n_estimators = 200, criterion = "entropy," class_weight = None, and verbose = 0, achieved an accuracy of 0.93, as shown in the previous table. Furthermore, Table 6 showed the testing outcomes for the NB algorithm.

Model	Parameter Settings	Accuracy
NB	(class_prior=None, fit_prior=False, alpha=1.0, force_alpha=True)	0.76

Table 6: Performance of NB

Table 6 showed the results of the NB method, configured with parameters class_prior = None, fit_prior = False, alpha = 1.0, and force_alpha = True, which achieved an accuracy of 0.76. Table 7 showed the testing outcomes for the KNN algorithm used in the study.

Model	Parameter Settings	Accuracy
KNN	(weights='uniform', metric_params=None, n_jobs=None, n_neighbors=3, metric='minkowski',)	0.81

Table 7: Performance of KNN

The results of the KNN method, configured with parameters weights = 'uniform', metric_params = None, n_jobs = None, n_neighbors = 3, and metric = 'minkowski', achieved an accuracy of 0.81, as shown in the previous table.

4.7 Design Widget Orange

Figure 7 showed the Orange widget design [Zulkifli et al., 2022] used to measure ROC analysis [Massoudi et al., 2021] on accuracy measurement results for predicting the survival of HF patient using NB, RF, KNN, SVM, and NN algorithm [By-nc-nd, 2018].

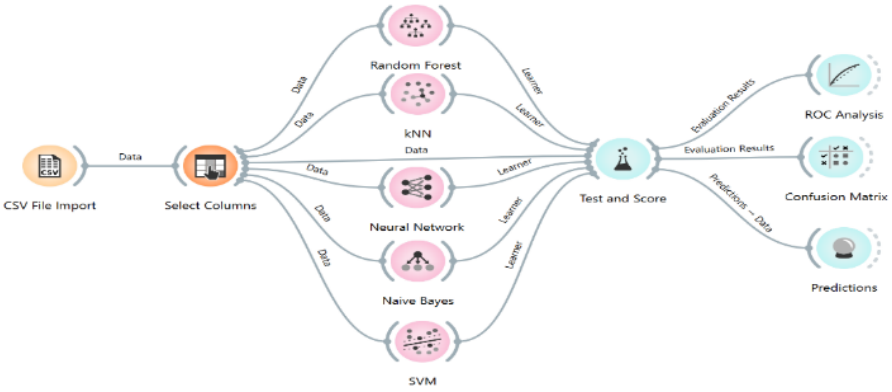

Figure 7: Orange widget design ROC analysis of the results of measuring the level of accuracy of predicting survival of HF patient using NB, RF, KNN, SVM, and NN algorithm.

4.8 Evaluation Confusion Matrix

Figure 8: Performance of the confusion matrix NN

The confusion matrix output from the NN algorithm was presented in the earlier figure. For the "survived" class, 17 instances were correctly classified, and 3 were misclassified. Similarly, for the "died" class, 7 instances were correctly classified, and 3 were misclassified, based on a total of 30 data points. Table 8 showed the evaluation results of the NN during the analysis [Ruuska et al., 2018].

	Precision	Recall	F-Measure	Class
	0.85	0.85	0.85	Survived
	0.70	0.70	0.70	Died
Weighted Avg.	0.80	0.80	0.80	

Table 8: Results of NN

Figure 9: Performance of the confusion matrix SVM

The confusion matrix output from the SVM algorithm was shown in the previous figure. For the "survived" class, 19 instances were correctly classified, and 5 were misclassified. Following the process, 5 instances were correctly classified, and 1 was misclassified for the "died" class, based on a total of 30 data points. Table 9 showed the evaluation results of the SVM during the study [Mustaqeem, 2021].

	Precision	Recall	F-Measure	Class
	0.79	0.95	0.86	Survived
	0.83	0.50	0.62	Died
Weighted Avg.	0.81	0.80	0.78	

Table 9: Performance Results of SVM

Figure 10: Performance of the confusion matrix RF

The confusion matrix output from the RF algorithm was shown in the previous figure. For the "survived" class, 19 instances were correctly classified, and 4 were misclassified. Similarly, 6 instances were correctly classified, and 1 was misclassified for the "died" class, based on a total of 30 data points. Table 10 showed the evaluation results of the RF during the analysis [Soe, 2018].

	Precision	Recall	F-Measure	Class
	0.83	0.95	0.88	Survived
	0.86	0.60	0.71	Died
Weighted Avg.	0.84	0.83	0.82	

Table 10: Performance Results of RF

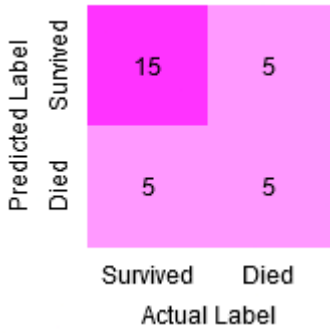

Figure 11: Performance of the confusion matrix NB

The confusion matrix output from the NB algorithm was shown in the earlier figure. For the "survived" class, 15 instances were correctly classified, and 5 were misclassified. Following the process, 5 instances were correctly classified, and 5 were misclassified for the "died" class, based on a total of 30 data points. Table 11 showed the evaluation results of NB [Zulkifli et al., 2024].

	Precision	Recall	F-Measure	Class
	0.75	0.75	0.75	Survived
	0.50	0.50	0.50	Died
Weighted Avg.	0.67	0.67	0.67	

Table 11: Performance Results of NB

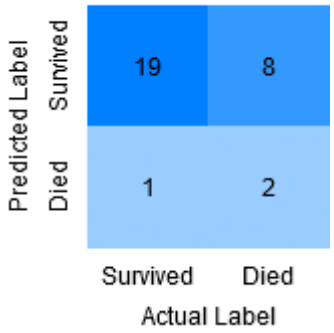

Figure 12: Performance of the confusion matrix KNN

The confusion matrix output from the KNN algorithm was shown in the subsequent figure. For the "survived" class, 19 instances were correctly classified, and 8 were misclassified. Similarly, 2 instances were correctly classified, where 1 was misclassified for the "died" class, based on a total of 30 data points. Table 12 showed the evaluation results of KNN during this study [Goyal, 2022b].

	Precision	Recall	F-Measure	Class
	0.70	0.95	0.81	Survived
	0.67	0.20	0.31	Died
Weighted Avg.	0.69	0.70	0.64	

Table 12: Performance Results of KNN

4.9 ROC Analysis

Receiver Operating Characteristic (ROC) curves were used to visualize, assess, and compare the performance of the NN, SVM, RF, NB, and KNN algorithm during the analysis. Figure 13 showed the performance curves of these algorithm, with the target classes "Survived" and "Died" represented by blue, purple, brown, green, as well as cyan.

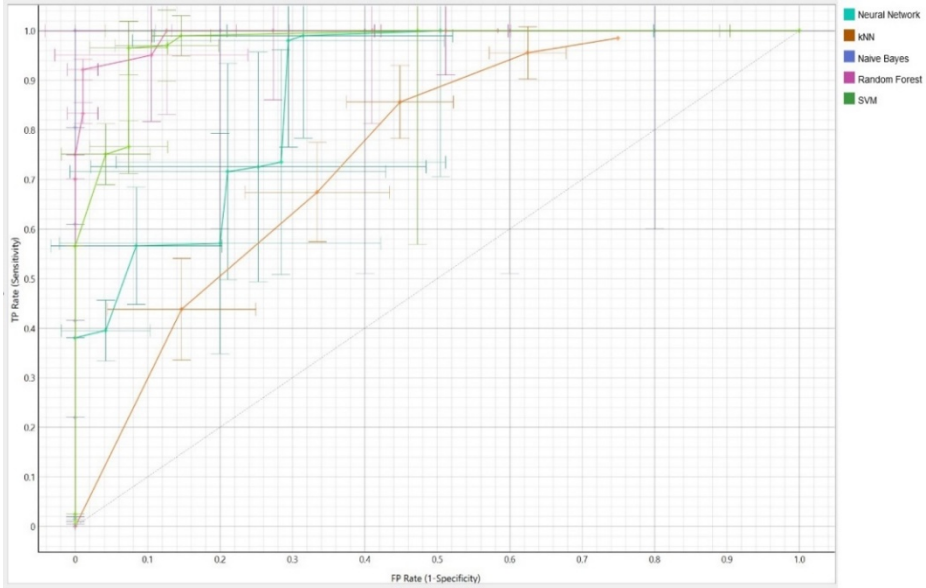

Figure 13: Performance curves of NB, RF, KNN, SVM, and NN algorithm

4.10 Comparison of NB, RF, KNN, SVM, and NN Algorithm for Predicting HF Patient Survival

Figure 14 showed the comparison of NB, RF, KNN, SVM, and NN algorithm for predicting the survival of HF patient.

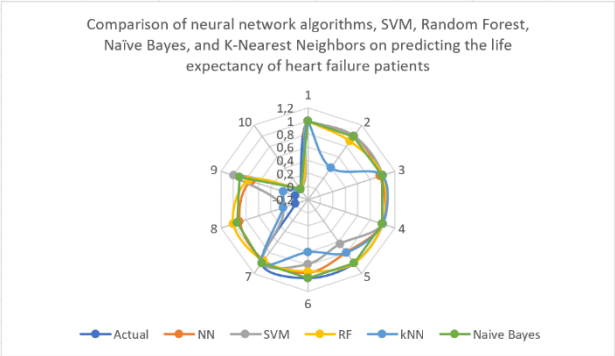

Figure 14: Comparison results of the performance of the accuracy level of HF patient survival prediction between NB, RF, KNN, SVM, and NN algorithm

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Patient-01	1	1	1	1	1	1	0.99	Died	0.99	Died	1	Died	0.99	Died	1	Died
Patient-02	1	1	1	1	1	0	0.99	Died	0.99	Died	0.9	Died	0.98	Died	0.4	Survived
Patient-03	1	1	1	1	1	1	0.96	Died	1	Died	1	Died	0.99	Died	1	Died
Patient-04	1	1	1	1	1	1	0.99	Died	0.99	Died	1	Died	0.99	Died	1	Died
Patient-05	1	1	1	1	1	1	0.8	Died	0.63	Died	1	Died	0.99	Died	0.8	Died
Patient-06	1	1	1	1	1	1	0.92	Died	0.78	Died	0.9	Died	0.99	Died	0.6	Survived
Patient-07	1	1	1	1	1	1	0.95	Died	0.99	Died	0.9	Died	0.99	Died	1	Died
Patient-15	0	0	0	0	0	1	0.89	Died	0.2	Survived	1	Died	0.93	Died	0.2	Survived
Patient-21	0	0	0	0	0	1	0.70	Died	0.99	Died	0.7	Died	0.9	Died	0.2	Survived
Patient-31	0	1	1	1	1	1	0	Survived	0	Survived	0	Survived	0	Survived	0	Survived

Note: 1=Patient Code, 2=Actual, 3=NN, 4=SVM, 5=RF, 6=NB, 7=KNN, 8= NN Numerical, 9= NN Validation Against Actual, 10=SVM Numerical, 11=SVM Validation Against Actual, 12=RF Numerical, 13=RF Validation Against Actual, 14=NB Numerical, 15=Validation of NB Against Actual, 16= KNN Numerical, 17= Validation of KNN Against Actual

Table 13: Validation results of survival prediction of HF patient with NB, RF, KNN, SVM, and NN algorithm

Many existing studies remained largely limited in scope and clinical applicability despite significant progress in ML applications for HF prediction. Most previous studies focused on evaluating a single algorithm using a limited number of clinical variables, such as ejection fraction or serum creatinine, without incorporating the model into practical tools for healthcare professionals. Moreover, current diagnostic methods, including imaging-based or invasive procedures such as angiography, were often expensive and inaccessible in low-resource healthcare settings. These limitations led to

low generalizability, poor usability, and limited support for real-time decision-making. To address the challenges, this study proposed a multi-algorithm web-based system that incorporated five ML algorithm, namely NB, RF, K-NN, SVM, and NN, in a single platform. The proposed system offered several advantages, including enabling comprehensive comparison of algorithm performance, providing an affordable and user-friendly interface for clinical use, as well as improving prediction reliability by incorporating 12 patient-related variables. Therefore, the motivation of this study was to bridge the gap between algorithmic development and clinical implementation by creating a practical, accurate, and accessible tool for predicting as well as monitoring HF patient survival.

5 Conclusions and future work

In conclusion, this study developed a web-based system for monitoring and predicting the survival of HF patient using five ML algorithm, namely NB, RF, KNN, SVM, and NN. NB achieved the highest accuracy of 88.21% among all the model. The originality of this work was in combining multi-algorithm in one software platform that could assist doctors and hospitals in diagnosis as well as prognosis. Different from previous studies that only focused on model development, this analysis produced a practical tool for clinical use based on 12 patient variables. However, the model was limited by the small dataset and the lack of incorporation with real-time hospital data. Future studies should use larger, multi-center data and include explainable AI methods to improve interpretability. Developing mobile or cloud-based versions could also make the system more accessible and efficient for healthcare providers.

Acknowledgements

The authors are grateful to the Ministry of Higher Education, Science, and Technology of the Republic of Indonesia for the support when conducting this study.

References

[Abdul-Samad et al., 2024] Abdul-Samad, K., Ma, S., Austin, D. E., Chong, A., Wang, C. X., Wang, X., Austin, P. C., Ross, H. J., Wang, B., & Lee, D. S. (2024). Comparison of machine learning and conventional statistical modeling for predicting readmission following acute heart failure hospitalization. American Heart Journal, 277, 93–103. https://doi.org/10.1016/j.ahj.2024.07.017

[Akdag et al., 2022] Akdag, S., Kuncan, F., & Kaya, Y. (2022). A new method for congestive heart failure and arrhythmia classification using downsampling local binary patterns with LSTM. Turkish Journal of Electrical Engineering and Computer Sciences, 30(6), 2145–2164. https://doi.org/10.55730/1300-0632.3930

[Akerman et al., 2025] Akerman, A. P., Al-Roub, N., Angell-James, C., Cassidy, M. A., Thompson, R., Bosque, L., Rainer, K., Hawkes, W., Piotrowska, H., Leeson, P., Woodward, G., Pellikka, P. A., Upton, R., & Strom, J. B. (2025). External validation of artificial intelligence for detection of heart failure with preserved ejection fraction. Nature Communications , 16(1), 1–12. https://doi.org/10.1038/s41467-025-58283-7

[AlDera & Othman, 2022] AlDera, S. A., & Othman, M. T. Ben. (2022). A Model for Classification and Diagnosis of Skin Disease using Machine Learning and Image Processing

Techniques. International Journal of Advanced Computer Science and Applications, 13(5), 252–259. https://doi.org/10.14569/ijacsa.2022.0130531

[Ali et al., 2023] Ali, S. A., Roy, N. R., & Raj, D. (2023). Software Defect Prediction using Machine Learning. Proceedings of the 17th INDIACom; 2023 10th International Conference on Computing for Sustainable Global Development, INDIACom 2023, Icoei, 639–642.

[Aydin et al., 2007] Aydin, I., Karaköse, M., & Akin, E. (2007). Artificial immune based support vector machine algorithm for fault diagnosis of induction motors. International Aegean Conference on Electrical Machines and Power Electronics and Electromotion ACEMP'07 and Electromotion'07 Joint Conference, 217–221. https://doi.org/10.1109/ACEMP.2007.4510505

[By-nc-nd, 2018] By-nc-nd, C. C. (2018). Zupan, Demsar: Introduction to Data Mining. May, 1–72.

[Byun et al., 2019] Byun, T., Sharma, V., Vijayakumar, A., & ... (2019). Input prioritization for testing neural networks. IEEE International Conference on Artificial Intelligence Testing.

[Chicco & Jurman, 2020] Chicco, D., & Jurman, G. (2020). Machine learning can predict survival of patient with heart failure from serum creatinine and ejection fraction alone. BMC Medical Informatics and Decision Making, 20(1), 1–16. https://doi.org/10.1186/s12911-020-1023-5

[Coşkun, 2022] Coşkun, C. (2022). Evaluation of Performance of Classification Algorithm in Prediction of Heart Failure Disease KalYetmezliği HastaliğininTahmiEdilmesinde SiniflandiriciAlgoritmalarinin PerformanslarinDeğerlendirilmesi. KSU J Eng Sci, 25(4), 1–4.

[Cross et al., 1995] Cross, S. S., Harrison, R. F., & Kennedy, R. L. (1995). Introduction to neural networks. In The Lancet (Vol. 346, Issue 8982). https://doi.org/10.1016/S0140-6736(95)91746-2

[Dritsas & Trigka, 2022] Dritsas, E., & Trigka, M. (2022). Data-Driven Machine-Learning Methods for Diabetes Risk Prediction. Sensors, 22(14). https://doi.org/10.3390/s22145304

[Esteban et al., 2019] Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., & ... (2019). fMRIPrep: a robust preprocessing pipeline for functional MRI. Nature ….

[Ghaffar Nia et al., 2023] Ghaffar Nia, N., Kaplanoglu, E., & Nasab, A. (2023). Evaluation of artificial intelligence techniques in disease diagnosis and prediction. Discover Artificial Intelligence, 3(1). https://doi.org/10.1007/s44163-023-00049-5

[Giri et al., 2021] Giri, B., Pandey, S., Shrestha, R., Pokharel, K., Ligler, F. S., & Neupane, B. B. (2021). Review of analytical performance of COVID-19 detection methods. Analytical and Bioanalytical Chemistry, 413(1), 35–48. https://doi.org/10.1007/s00216-020-02889-x

[Goyal, 2022a] Goyal, S. (2022a). Handling Class-Imbalance with KNN (Neighbourhood) Under-Sampling for Software Defect Prediction. In Artificial Intelligence Review (Vol. 55, Issue 3). Springer Netherlands. https://doi.org/10.1007/s10462-021-10044-w

[Goyal, 2022b] Goyal, S. (2022b). Handling Class-Imbalance with KNN (Neighbourhood) Under-Sampling for Software Defect Prediction. Artificial Intelligence Review, 55(3), 2023–2064. https://doi.org/10.1007/s10462-021-10044-w

[Huang et al., 2023] Huang, J., Yeung, A. M., Armstrong, D. G., Battarbee, A. N., Cuadros, J., Espinoza, J. C., Kleinberg, S., Mathioudakis, N., Swerdlow, M. A., & Klonoff, D. C. (2023). Artificial Intelligence for Predicting and Diagnosing Complications of Diabetes. Journal of Diabetes Science and Technology, 17(1), 224–238. https://doi.org/10.1177/19322968221124583

[Ismael & Şengür, 2021] Ismael, A. M., & Şengür, A. (2021). Deep learning methods for COVID-19 detection based on chest X-ray images. Expert Systems with Applications, 164(March 2020). https://doi.org/10.1016/j.eswa.2020.114054

[Kerr & Lagen, 2021] Kerr, W. J., & Lagen, J. B. (2021). Cardiovascular diseases. In Medicine of the year (Vol. 4, Issue 2).

[Mallikharjuna Rao et al., 2023] Mallikharjuna Rao, K., Saikrishna, G., & Supriya, K. (2023). Data preprocessing techniques: emergence and selection towards machine learning model - a practical review using HPA dataset. Multimedia Tools and Applications, 82(24), 37177–37196. https://doi.org/10.1007/s11042-023-15087-5

[Mandal et al., 2021] Mandal, B., Okeukwu, A., & Theis, Y. (2021). Masked Face Recognition using ResNet-50. http://arxiv.org/abs/2104.08997

[Massoudi et al., 2021] Massoudi, M., Jain, N. K., & Bansal, P. (2021). Software defect prediction using dimensionality reduction and deep learning. Proceedings of the 3rd International Conference on Intelligent Communication Technologies and Virtual Mobile Networks, ICICV 2021, Icicv, 884–893. https://doi.org/10.1109/ICICV50876.2021.9388622

[Meeradevi et al., 2025] Meeradevi, T., Sasikala, S., Murali, L., Manikandan, N., & Ramaswamy, K. (2025). Lung cancer detection with machine learning classifiers with multi-attribute decision-making system and deep learning model. Scientific Reports, 15(1), 1–19. https://doi.org/10.1038/s41598-025-88188-w

[Mustaqeem, 2021] Mustaqeem, M. (2021). Principal component based support vector machine (PC-SVM): a hybrid technique for software defect detection. Cluster Computing, 24(3), 2581–2595. https://doi.org/10.1007/s10586-021-03282-8

[Niu, 2018] Niu, Y. (2018). Adaptive two-SVM multi-objective cuckoo search algorithm for software defect prediction. International Journal of Computing Science and Mathematics, 9(6), 547–554. https://doi.org/10.1504/IJCSM.2018.096327

[Petry, 2020] Petry, K. L. (2020). Model-based testing of software product lines: Mapping study and study roadmap. Journal of Systems and Software, 167. https://doi.org/10.1016/j.jss.2020.110608

[Ruuska et al., 2018] Ruuska, S., Hämäläinen, W., Kajava, S., Mughal, M., Matilainen, P., & Mononen, J. (2018). Evaluation of the confusion matrix method in the validation of an automated system for measuring feeding behaviour of cattle. Behavioural Processes, 148, 56–62. https://doi.org/10.1016/j.beproc.2018.01.004

[Saboor et al., 2022] Saboor, A., Usman, M., Ali, S., Samad, A., Abrar, M. F., & Ullah, N. (2022). A Method for Improving Prediction of Human Heart Disease Using Machine Learning Algorithm. Mobile Information Systems, 2022. https://doi.org/10.1155/2022/1410169

[Singh & Verma, 2018] Singh, P., & Verma, S. (2018). Multi-classifier model for software fault prediction. International Arab Journal of Information Technology, 15(5), 912–919.

[Soe, 2018] Soe, Y. (2018). Software defect prediction using random forest algorithm. In Proceedings - 12th SEATUC Symposium, SEATUC 2018. https://doi.org/10.1109/SEATUC.2018.8788881

[Tripoliti et al., 2017] Tripoliti, E. E., Papadopoulos, T. G., Karanasiou, G. S., Naka, K. K., & Fotiadis, D. I. (2017). Heart Failure: Diagnosis, Severity Estimation and Prediction of Adverse Events Through Machine Learning Techniques. Computational and Structural Biotechnology Journal, 15, 26–47. https://doi.org/10.1016/j.csbj.2016.11.001

[Umer et al., 2022] Umer, M., Sadiq, S., Karamti, H., Karamti, W., Majeed, R., & Nappi, M. (2022). IoT Based Smart Monitoring of Patient' with Acute Heart Failure. Sensors, 22(7), 1–18. https://doi.org/10.3390/s22072431

[Varshni et al., 2019] Varshni, D., Thakral, K., Agarwal, L., Nijhawan, R., & Mittal, A. (2019). Pneumonia Detection Using CNN based Feature Extraction. Proceedings of 2019 3rd IEEE International Conference on Electrical, Computer and Communication Technologies, ICECCT 2019. https://doi.org/10.1109/ICECCT.2019.8869364

[H. Wang et al., 2021] Wang, H., Czerminski, R., & Jamieson, A. C. (2021). Neural Networks and Deep Learning. The Machine Age of Customer Insight, 91–101. https://doi.org/10.1108/978-1-83909-694-520211010

[K. Wang et al., 2021] Wang, K., Tian, J., Zheng, C., Yang, H., Ren, J., Liu, Y., Han, Q., & Zhang, Y. (2021). Interpretable prediction of 3-year all-cause mortality in patient with heart failure caused by coronary heart disease based on machine learning and SHAP. Computers in Biology and Medicine, 137, 104813. https://doi.org/10.1016/j.compbiomed.2021.104813

[Xie et al., 2025] Xie, Y., Zhang, L., Sun, W., Zhu, Y., Zhang, Z., Chen, L., Xie, M., & Zhang, L. (2025). Artificial Intelligence in Diagnosis of Heart Failure. Journal of the American Heart Association, 14(8), 1–14. https://doi.org/10.1161/JAHA.124.039511

[Xu et al., 2021] Xu, Y., Peng, C., Li, M., Li, Y., & Du, S. (2021). Pyramid Feature Attention Network for Monocular Depth Prediction. Proceedings - IEEE International Conference on Multimedia and Expo. https://doi.org/10.1109/ICME51207.2021.9428446

[Zulkifli et al., 2022] Zulkifli, Gaol, F. L., Trisetyarso, A., & Budiharto, W. (2022). Software Testing Model by Measuring the Level of Accuracy Fault Output Using Neural Network Algorithm. Proceedings of the 2022 IEEE 7th International Conference on Information Technology and Digital Applications, ICITDA 2022, 1–6. https://doi.org/10.1109/ICITDA55840.2022.9971274

[Zulkifli et al., 2024] Zulkifli, Z., Makkiyah, F. A., Antoni, D., Fitriana, F., Jamaan, T., & Taufik, A. (2024). Multi-Algorithm to Measure the Accuracy Level of Diabetes Status Prediction. Journal of Applied Data Sciences, 5(2), 736–746. https://doi.org/10.47738/jads.v5i2.250