Jurnal ELTIKOM:
Jurnal Teknik Elektro, Teknologi Informasi dan Komputer

Vol. 7, No. 2, December 2023, page. 116-124
ISSN 2598-3245 (Print), ISSN 2598-3288 (Online)
DOI: http://doi.org/10.31961/eltikom.v7i2.822
Available online at http://eltikom.poliban.ac.id

RANDOM STATE INITIALIZED LOGISTIC REGRESSION FOR IMPROVED HEART ATTACK PREDICTION

Kevyn Alifian Hernanda Wibowo[1*], Salma Aprilia Huda Putri[1], Jumanto[1], Much Aziz Muslim[2]
[1)] Department of Computer Science, Universitas Negeri Semarang, Indonesia
[2)] Faculty of Technology Management, University Tun Hussein Onn Malaysia, Malaysia
e-mail: kevinalifian2802@students.unnes.ac.id, salmaahputri@students.unnes.ac.id, jumanto@mail.unnes.ac.id, a212muslim@mail.unnes.ac.id

Received: 7 July 2023 – Revised: 28 September 2023 – Accepted: 5 October 2023

ABSTRACT

One of the primary causes of death in Indonesia is heart attacks. Therefore, an effective method of pre-diction is required to determine whether a patient is experiencing a heart attack. One efficient approach is to use machine learning models. However, it is still rare to find machine learning models that have good performance in predicting heart attacks. This study aims to develop a machine learning model on Logistic Regression algorithm in predicting heart attack. Logistic Regression is one of the machine learning meth-ods that can be used to study the relationship between a binary response variable [0,1] and a set of pre-dictor variables, and can be used directly to calculate probabilities. In this study, a random state is ini-tialized in the Logistic Regression model in order to stabilize the training of the machine learning model and increase the precision of the proposed method. The results of this study show that the proposed model can be a method that has good performance in predicting heart attack disease.

Keywords: *heart attack, logistic regression, machine learning, prediction, random state.*

I. INTRODUCTION

DISEASE prediction has become very important and relevant in human life, and several studies have been conducted to develop and improve disease prediction systems [1]–[3]. The heart is the most important part of the human body responsible for pumping oxygen-rich blood to all parts of the body through a network of arteries and veins [4]. According to the World Health Organization (WHO), heart attacks are caused by blockage of arteries due to fat, and symptoms include chest pain and shortness of breath. Heart disease is the leading cause of death in Indonesia because there are factors that can trigger a heart attack, including narrowing of blood vessels, infections caused by bacteria and viruses, unhealthy lifestyles, disorders of heart valves, and the influence of the use of certain drugs. The heart has four chambers, namely two porches at the top and two chambers at the bottom. The function of the muscle wall between the left and right chambers (septum) is to separate oxygen-rich blood from deoxygenated blood [5]–[8].

According to the World Heart Federation, in 2014, Southeast Asia recorded approximately 1.8 million deaths due to heart failure. Meanwhile, more than 883,447 cases of heart failure have been recorded in Indonesia, with most patients aged 55 to 64. Approximately 45% of all fatalities in Indonesia are caused by cardiovascular disease [9]. A number of studies have been conducted utilizing machine learning models. However, there is currently no machine learning model that performs well in predicting heart attacks, such as the Multi-Layer Perceptron for Enhanced Brownian Motion based on the Dragonfly Algorithm [10],[11] and research using the cluster-based bidirectional Long Short Term Memory Network algorithm [12],[13].

Other studies used the Grey Wolf Horse Herd machine learning algorithm method based on Shepard Convolutional Neural Network optimization [14], a pre-trained Deep Neural Network for feature extraction, Principal Component Analysis for dimension reduction, and Logistic Regression for

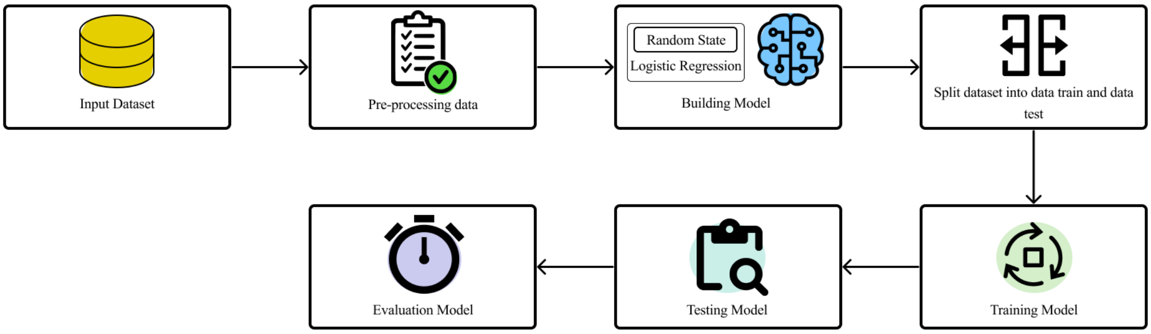

Figure 1. Flowchart of the proposed method

prediction [15]. Some studies also used the X^2 statistical model when optimally configuring the Deep Neural Network [3], a stacking classifier method to predict heart attack classification [16]. Although research employing a trained Deep Neural Network for feature extraction, Principal Component Analysis for dimension reduction, and Logistic Regression for prediction has demonstrated an accuracy of 93.3%, the proposed model employs Logistic Regression for prediction in a manner similar to the aforementioned research. However, further research is needed to obtain better performance and more accurate prediction values in predicting heart attacks. This study proposes a random state initialized in the Logistic Regression method. To solve the problem of heart attack prediction, this study aims to utilize the initialized random state in the Logistic Regression method to stabilize and improve the performance in predicting heart attacks.

II. RESEARCH METHOD

Logistic Regression is used in studying the relationship between binary response variables {0,1} and a set of predictor variables. Therefore, predictable values in Logistic Regression are between "yes or no" and they can be directly used as probabilities [17]. A flowchart for the proposed method for Heart Attack Prediction can be seen in Figure 1 and further explanation is explained afterwards.

A. Input Dataset

Heart Attack Analysis and Prediction Dataset is the dataset used in this study. It is accessed via the Kaggle Repository at the following link: https://www.kaggle.com/datasets/rashikrahmanpritom/heart-attack-analysis-prediction-dataset and consists of 2 CSV files. The heart.csv file contains 14 attributes, namely age, sex, cp, trtbps, chol, fbs, restecg, thalachh, exng, oldpeak, slp, caa, thall, and output and the o2saturaion file which contains only 1 column. Furthermore, the preprocessing step is performed.

Pre-processing data is the initial stage of data before data processing is performed [18]. The union of 2 files .csv is required in the data pre-processing stage. In the subsequent phase, column names are modified to make them more understandable, such as "cp" to chest_pain, "trtbps" to blood_pressure, "chol" to cholesterol, "fbs" to fasting_blood_sugar, "thalachh" to max_heart_rate, "exng" to angina, "caa" to n_vessels, "output" to heart_attack. In addition, the binary numbers [0, 1] in the sex column are changed to (0 to Female and 1 to Male).

B. Building Model

To build a Logistic Regression model, it is necessary to include the LogisticRegression library from sklearn.linear_model. Other libraries such as numpy, pandas, train_test_split, and accuracy_score are also needed in the process of building Machine Learning models to predict heart attacks. The concept of Logistic Regression is to study the relationship between a set of variables as input and a binary variable [0,1] as output and can be directly used as a probability. Figure 1 shows the visualized concept of Logistic Regression.

Logistic regression models are able to produce high accuracy in making predictions and also have the ability to use their input data to produce probabilities [19]. Logistic Regression is a popular classification algorithm that models the relationship between a dependent variable and one or more independent variables using a logistic function. Its ability to predict probabilities based on input features makes it a good choice for predicting the likelihood of an event occurring. Equation (1) shows the calculation on Logistic Regression [20].

$$P = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n)}} \tag{1}$$

Where P the probability of the event occuring, $\beta_0, \beta_1, \ldots, \beta_n$ are the coefficients of the model, and $x_1, \ldots, x_n$ are the predictor variables [20].

C. Split dataset into data training and data test

The dataset used contains 15 columns that each have 304 rows. A 65-time random state is used to divide the dataset into 80 percent training data and 20 percent test data. As a result of experimentation, these 65 random states can significantly improve the accuracy of the model used in this study. For each state sampled from heart attack dataset, a new random state can be generated using (2) [21].

$$h' = h + \frac{c}{a}\ (h_{max} - h_{min}) \tag{2}$$

Equation (2) h_{min} and h_{max} are the minimum and maximum specific enthalpies of the heart attack dataset, respectively. Denotation h is the specific enthalpy of the sampled heart attack state, h' is the randomly generated specific enthalpy value. This will cause the random data to deviate greatly from heart attack data. To solve this problem, (3) and (4) is created [21].

$$y'_j = y_j^{(1+\frac{c}{b})} \tag{3}$$

$$y''_j = \frac{y'_j}{\sum y'} \tag{4}$$

In Equation (3), y_j is the concentration of the jth species of the heart attack data, y'_j is the randomly generated species concentration, c is a random number between -1 to 1, while b is the coefficient. Equations (2) to (4) are used to create random states, and then the corresponding attributes are calculated. If all constraints are met, then the new state is accepted. Otherwise, it is rejected and the process is repeated until the data is accepted or the number of rejections reaches 65. In (2) and (3), the coefficients a and b, respectively, are determined to satisfy these constraints. This is because very large coefficients make the random data unrealistic and increase the difficulty in satisfying all the constraints, while small coefficients make the random data very similar to the heart attack dataset [21].

D. Training Model

Variables X and Y are needed to train an already-built model to predict heart attacks using the Logistic Regression Method. The X variables used are sex, age, chest_pain, blood_pressure, cholestoral, fasting_blood_sugar, restecg, max_heart_rate, angina, oldpeak, slp, n_vessels, thall, while the Y variable used is heart_attack.

E. Testing Model

Variables X and Y are also required to assess the model's ability to predict heart attacks using data that the model has never seen before. To determine whether the model is successful, the accuracy of the model in predicting heart attacks must be displayed using the accuracy_score library.

F. Evaluation Model

A confusion Matrix can present predictions and actual conditions on data generated by machine learning algorithms. In the Confusion Matrix, True Positives (*TP*), True Negatives (*TN*), False Positives (*FP*), and False Negatives (*FN*) are classes commonly used in the confusion matrix which include the number of true positive classes, the number of false positive classes, true negative classes, and false negative classes in the data. Equation (5) is used to calculate accuracy. Accuracy is determined by the proportion of correctly assigned samples and the overall sample size for the test dataset [22].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

Recall is a measure of the ratio of correct positive predictions to all correct positive data, i.e., *TP* and *FN* [22]. Equation (6) is used to calculate recall.

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

Equation (7) is used to calculate the *F1 score*, *F1 score* or the average value of recall (the ability to find all positive data samples) and precision (the ability not to label a negative as a positive class) [22].

$$f1\ score = \frac{2 \times precision \times recall}{precision + recall} \tag{7}$$

III. RESULT AND DISCUSSION

A. Data Analysis Process

This study used the Kaggle Heart Attack dataset containing observational data to evaluate the performance of the Logistic Regression method. This dataset consists of 304 samples that are used as sample data in data analysis. Detailed information about the Heart Attack dataset can be seen in Table 1.

This dataset contains 304 samples covering 14 attributes that are used in data processing. In the dataset, there are two labels that indicate the probability of having a heart attack, namely 0 which indicates a lower probability, and 1 which indicates a higher probability. The following Table 2 is an explanation for each attribute in the dataset used.

B. Data Pre-processing

Death from a heart attack often occurs after a stroke and can occur without any specific age limit. This study uses the Logistic Regression method to process large amounts of data, to obtain accurate results. Table 3 describes the definition of the Dataset Header.

C. Attribute Data Type Check

After completing the data processing, the process continues by verifying the type of data to be used as an attribute, to determine whether the attribute value is an integer or an array of strings as shown in Table 4.

TABLE 1
DATASET HEART ATTACK

	age	sex	cp	trtbps	chol	fbs	restecg	thalachh	exng	oldpeak	slp	caa	thall	output	o2Saturation
0	63	1	3	145	233	1	0	150	0	2.3	0	0	1	1	98.6
1	37	1	2	130	250	0	1	187	0	3.5	0	0	2	1	98.6
2	41	0	1	130	204	0	0	172	0	1.4	2	0	2	1	98.6
3	56	1	1	120	236	0	1	178	0	0.8	2	0	2	1	98.6
4	57	0	0	120	354	0	1	163	1	0.6	2	0	2	1	98.1
5	57	1	0	140	192	0	1	148	0	0.4	1	0	1	1	97.5
6	56	0	1	140	294	0	0	153	0	1.3	1	0	2	1	97.5

TABLE 2
HEART ATTACK DATASET ATTRIBUTES

No	Attributes	Detail
1	age	Age of the patient
2	sex	Sex of the patient
3	cp	Chess pain type
4	trtbps	Resting blood pressure
5	chol	Cholesterol in mg/dl
6	fbs	Fasting blood sugar
7	restecg	Resting electrocardiographic results
8	thalachh	Maximum heart rate achieved
9	exng	Exercise induced angina
10	oldpeak	Previous peak
11	slp	The slope of the peak exercise
12	caa	Number of major vessels
13	thall	Thal rate
14	output	Target variable
15	o2saturation	Saturation level

TABLE 3
DEFINING DATASET HEADER

	count	mean	std	min	25%	50%	75%	max
age	303.0	54.366337	9.082101	29.0	47.5	55.0	61.0	77.0
blood_pressure	303.0	131.623762	17.538143	94.0	120.0	130.0	140.0	200.0
cholestoral	303.0	246.264026	51.830751	126.0	211.0	240.0	274.5	564.0
max_heart_rate	303.0	149.646865	22.905161	71.0	133.5	153.0	166.0	202.0
o2saturation	303.0	97.484488	0.352649	96.5	97.5	97.5	97.5	98.6

TABLE 4
DATA TYPE CHECK

	Column	Non-Null Count	Dtype
0	age	303 non-null	int64
1	sex	303 non-null	int64
2	cp	303 non-null	int64
3	trtbps	303 non-null	int64
4	chol	303 non-null	int64
5	fbs	303 non-null	int64
6	restecg	303 non-null	int64
7	thalachh	303 non-null	int64
8	exng	303 non-null	int64
9	oldpeak	303 non-null	float64
10	slp	303 non-null	int64
11	caa	303 non-null	int64
12	thall	303 non-null	int64
13	output	303 non-null	int64
14	o2saturation	303 non-null	float64

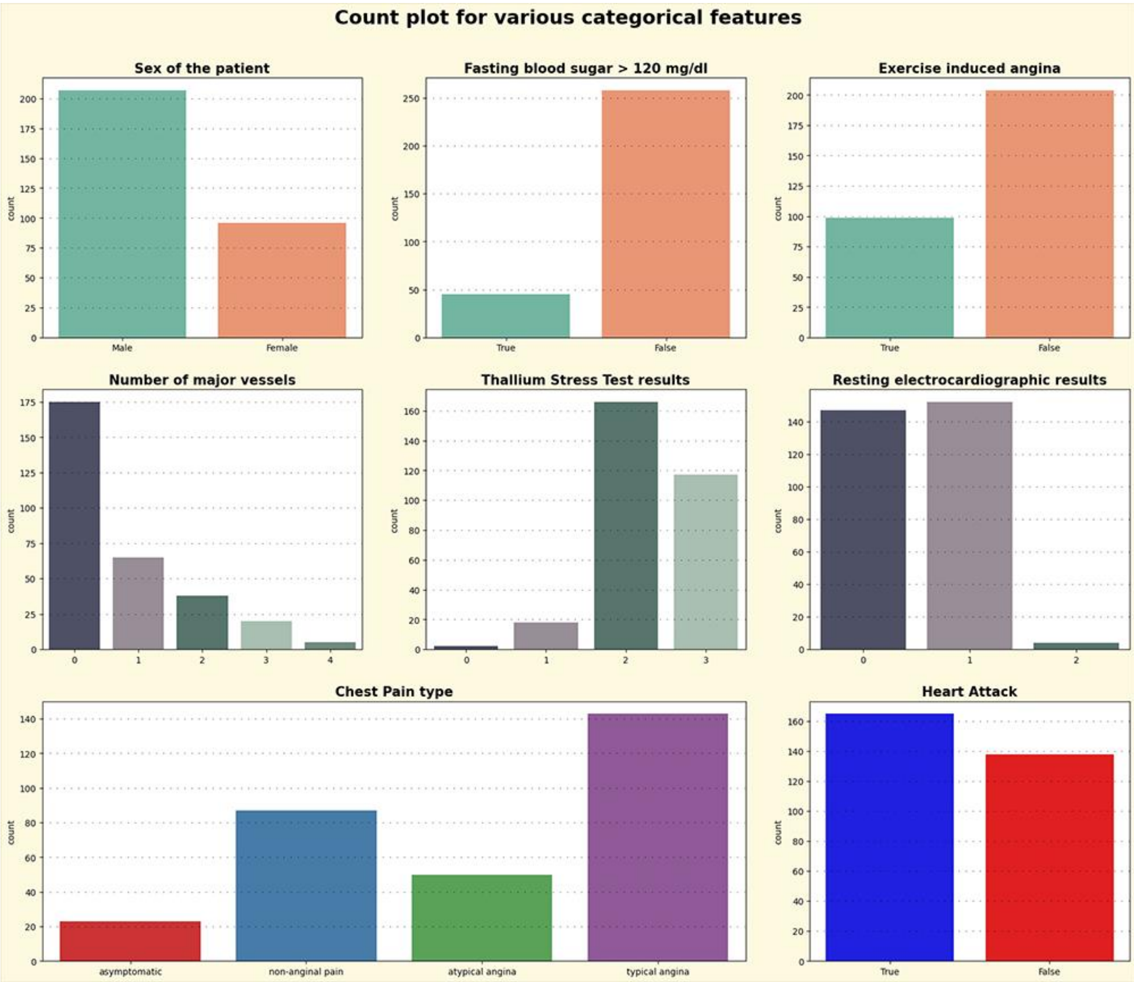

Figure 2. Count plot for various categorial features

After the data type checking process is carried out, it is then continued by separating columns in the form of categories and continuous. This separation is used to facilitate the data processing process carried out.

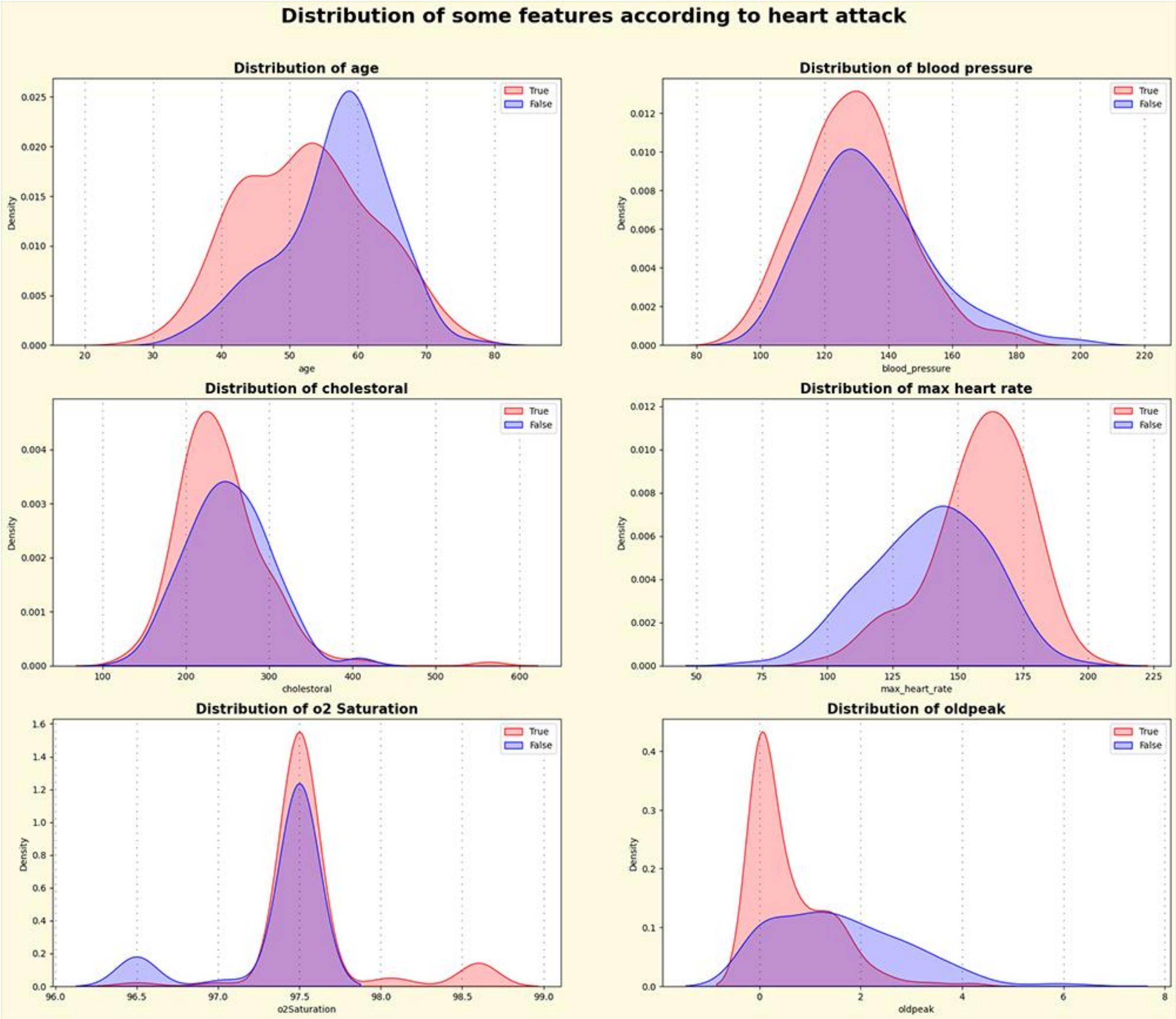

Figure 3. Distribution of some features according to heart attack

D. Data Visualization Process

The purpose of data visualization is to transform a data set into a more complex or simple format so that it is easier to understand and analyze. In general, these data visualizations usually take the form of histograms or line graphs. Figure 2 visualizes graphs of patient gender, fasting blood sugar, blood pressure, Exercise Induced angina, number of major blood vessels, Thallium Stress Test results, resting electrocardiography results, chest pain type, and Heart Attack. Based on Figure 2, it can be determined in the patient's Gender graph that male patients experience more heart attacks than female patients. Then, based on the graph of fasting blood sugar, it can be determined that only a small number of patients have blood sugar levels greater than 120 mg/dL. The Exercise Induced angina graph reveals that a considerable number of patients experience general heart complaints during exercise. On the Resting electrocardiography results graph, patients with a result of 0 indicate that the patient is normal, patients with ST-T wave abnormalities are marked with a result of 1, and patients with possible or definite left ventricular hypertrophy according to the Estes criteria are marked with a result of 2 that is smaller than the previous Resting electrocardiography type. The Chest Pain type graph reveals that patients experience angina more frequently than other forms of Chest Pain. Furthermore, the Heart attack graph reveals that patients are reported to have experienced a greater number of heart attacks than those who did not experience heart attacks.

Figure 3 visualizes age distribution, blood pressure distribution, cholestoral distribution, maximum heart rate distribution, O2 saturation distribution, oldpeak distribution. Based on Figure 3, patients who are declared to have a heart attack are marked with true and patients who are declared not to have a heart attack are marked with false. It can be defined that in the Distribution of Age, patients who have a heart attack have an average age of 50 to 60 years. Then, in the blood pressure distribution, heart attack patients had blood pressure ranging from 120 to 140 mmHg. The average cholesterol level of heart attack patients is between 200 and 300 mg/dL, as shown in the Cholesterol Distribution. Heart attack patients averaged a maximum heart rate between 150 and 175 beats per minute. The Distribution of o2 Saturation reveals that patients who had a heart attack had an average blood oxygen level of 97.5%. Oldpeak is the level of depression experienced during rest; according to the Distribution of oldpeak, heart attack patients with oldpeak levels ranging from 0 to 1 had experienced depression during rest.

E. Data Processing

Processing raw data in a more structured format is done because raw data generally does not have a consistent format. It aims to transform data into a form that is easier to understand and can be used as a source of information through a set of data that is ready for further processing [23].

F. Result

This study is evaluated using the Heart Attack Analysis and Prediction Dataset, which consists of two CSV files containing 14 and 1 attributes, respectively. In the preprocessing phase, the 2 CSV files are unified and renamed for each attribute to make them easier to comprehend. After constructing a model with Logistic Regression, the dataset is divided into training data and test data with a ratio of 80:20, meaning that 80% of the dataset consists of training data and 20% of the dataset consists of test data.

TABLE 5
PERFORMANCES REPORT

Random State	Accuracy	Recall	F1-score
0	0.85	0.88	0.87
24	0.87	0.94	0.88
42	0.89	0.91	0.89
50	0.79	0.89	0.83
65	0.95	0.97	0.95

TABLE 6
COMPARISON OF PROPOSED MODEL WITH OTHER MODELS

Author	Model	Accuracy
V. Chang et al. [24]	Random Forest	83%
A. A. Almazroi et al. [25]	CNN	83%
A. A. Hussein [26]	K-Means	84.74%
M. Ozcan et al. [27]	Classification and Regression Tree (CART)	87%
M. Gjoreski et al. [28]	Fully Connected Neural Network (FCNN)	93.2%
D. Hassan et al. [15]	Logistic Regression	93.3%
S. A. Ali et al. [29]	Optimally configured and improved deep belief network (OCI-DBN)	94.61%
Proposed model	Logistic Regression	95%

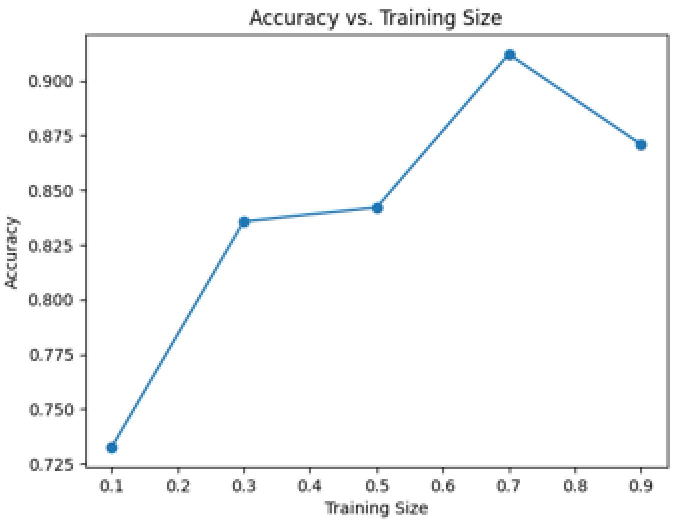

Figure 4. Logistic Regression algorithm accuracy graph

This determination is based on previous experiments and the initialization of the state to a random value of 65. Next, the training model is executed using training data that has been divided in accordance with a predetermined ratio, with 14 attributes other than attributes named heart_attack serving as variable X and 1 attribute named heart_attack serving as variable Y. After that, testing the model with test data using the same X variables and Y variables as in the previous step. Figure 4 depicts a visualization created to compare the training model process to the testing model process.

The results of model evaluation testing using Confusion Matrix which produces Accuracy, Recall and F1-Score values are shown in Table 5. Based on Table 5, initializing the random state with a value of 0 yields an accuracy of 85 percent. A random state up to 24 yields an 87% accuracy. A random number up to 42 produces an accuracy of 89%. A random state as high as 50 yields an accuracy of 7. According to the results of a comparison of random state initialization in the logistic regression model, random state 65 demonstrates the greatest performance, with a high degree of accuracy of 95%. Therefore, the results of this study indicate that random state initialization of 65 is the right solution to improve performance in the logistic regression model used in this study.

G. Discussion

In an effort to improve the quality of predictions, a novel model that compares to existing models is developed. It analyzes the predictive accuracy of both models and presents comparative findings, making a significant contribution to our comprehension of the benefits and drawbacks of each model. Table 6 demonstrates that the constructed model is more accurate than previous models.

IV. CONCLUSION

On the basis of the test results and analysis of random state initialization in the logistic regression model, it is possible to conclude that, in terms of accuracy, the model is highly accurate. The predictive accuracy of this model is 95%. Despite the fact that this model has the highest predictive accuracy for heart attacks based on the used attributes, it is crucial to conduct a thorough analysis and conduct additional validation. This model may be one of the most efficient methods for identifying patients at high risk. Nevertheless, it is essential to recognize the limitations of predicting complex factors and interactions that are not accounted for by the model. In order to predict cardiac attacks with greater accuracy, it is anticipated that future research will employ additional machine learning methods.

REFERENCES

[1] M. S. Amin, Y. K. Chiam, and K. D. Varathan, "Identification of significant features and data mining techniques in predicting heart disease," *Telemat. Informatics*, vol. 36, pp. 82–93, 2019.

[2] R. P. Cherian, N. Thomas, and S. Venkitachalam, "Weight optimized neural network for heart disease prediction using hybrid lion plus particle swarm algorithm," *J. Biomed. Inform.*, vol. 110, p. 103543, 2020.

[3] L. Ali, A. Rahman, A. Khan, M. Zhou, A. Javeed, and J. A. Khan, "An automated diagnostic system for heart disease prediction based on X^2 statistical model and optimally configured deep neural network," *Ieee Access*, vol. 7, pp. 34938–34945, 2019.

[4] P. Rani, R. Kumar, N. M. O. S. Ahmed, and A. Jain, "A decision support system for heart disease prediction based upon machine learning," *J. Reliab. Intell. Environ.*, vol. 7, no. 3, pp. 263–275, 2021.

[5] H. W. Dhany, "Performa Algoritma K-Nearest Neighbour dalam Memprediksi Penyakit Jantung," in *Seminar Nasional Informatika (SENATIKA)*, 2021, pp. 176–179.

[6] E. Evangelista, "A hybrid machine learning framework for predicting students' performance in virtual learning environment," *Int. J. Emerg. Technol. Learn.*, vol. 16, no. 24, pp. 255–272, 2021.

[7] R. Annisa, "Analisis Komparasi Algoritma Klasifikasi Data Mining Untuk Prediksi Penderita Penyakit Jantung," *JTIK (Jurnal Tek. Inform. Kaputama)*, vol. 3, no. 1, pp. 22–28, 2019.

[8] P. D. Putra and D. P. Rini, "Prediksi Penyakit Jantung dengan Algoritma Klasifikasi," *Pros. Annu. Res. Semin*, vol. 5, no. 1, pp. 978–979, 2019.

[9] P. Subarkah, W. R. Damayanti, and R. A. Permana, "Comparison of correlated algorithm accuracy Naive Bayes Classifier and Naive Bayes Classifier for heart failure classification," *Ilk. J. Ilm.*, vol. 14, no. 2, pp. 120–125, 2022.

[10] D. Deepika and N. Balaji, "Effective heart disease prediction using novel MLP-EBMDA approach," *Biomed. Signal Process. Control*, vol. 72, p. 103318, 2022.

[11] A. M. Hemeida, S. Alkhalaf, A. Mady, E. A. Mahmoud, M. E. Hussein, and A. M. B. Eldin, "Implementation of nature-inspired optimization algorithms in some data mining tasks," *Ain Shams Eng. J.*, vol. 11, no. 2, pp. 309–318, 2020.

[12] P. Dileep *et al.*, "An automatic heart disease prediction using cluster-based bi-directional LSTM (C-BiLSTM) algorithm," *Neural Comput. Appl.*, vol. 35, no. 10, pp. 7253–7266, 2023.

[13] T. S. Pooja and P. Shrinivasacharya, "Evaluating neural networks using Bi-Directional LSTM for network IDS (intrusion detection systems) in cyber security," *Glob. Transitions Proc.*, vol. 2, no. 2, pp. 448–454, 2021.

[14] P. Kanchanamala, A. S. Alphonse, and P. V. B. Reddy, "Heart disease prediction using hybrid optimization enabled deep learning network with spark architecture," *Biomed. Signal Process. Control*, vol. 84, p. 104707, 2023.

[15] D. Hassan, H. I. Hussein, and M. M. Hassan, "Heart disease prediction based on pre-trained deep neural networks combined with principal component analysis," *Biomed. Signal Process. Control*, vol. 79, p. 104019, 2023.

[16] S. Mohapatra, S. Maneesha, P. K. Patra, and S. Mohanty, “Heart Diseases Prediction based on Stacking Classifiers Model,” *Procedia Comput. Sci.*, vol. 218, pp. 1621–1630, 2023.
[17] K. Kannan and A. Menaga, “Risk factor prediction by naive bayes classifier, logistic regression models, various classification and regression machine learning techniques,” *Proc. Natl. Acad. Sci. India Sect. B Biol. Sci.*, vol. 92, no. 1, pp. 63–79, 2022.
[18] V. A. Gunawan, I. I. Fitriani, and L. S. A. Putra, “Klasifikasi Rambu Lalu Lintas Menggunakan Ekstraksi Ciri Wavelet Dan Jarak Euclidean,” *J. ELTIKOM J. Tek. Elektro, Teknol. Inf. dan Komput.*, vol. 3, no. 1, pp. 26–35, 2019.
[19] M. Fang and D. T. N. Huy, “Building a cross-border e-commerce talent training platform based on logistic regression model,” *J. High Technol. Manag. Res.*, vol. 34, no. 2, p. 100473, 2023.
[20] S. Hafeez, S. S. Alotaibi, A. Alazeb, N. Al Mudawi, and W. Kim, “Multi-sensor-based Action Monitoring and Recognition via Hybrid Descriptors and Logistic Regression,” *IEEE Access*, 2023.
[21] T. Ding, T. Readshaw, S. Rigopoulos, and W. P. Jones, “Machine learning tabulation of thermochemistry in turbulent combustion: An approach based on hybrid flamelet/random data and multiple multilayer perceptrons,” *Combust. Flame*, vol. 231, p. 111493, 2021.
[22] M. A. Muslim, T. L. Nikmah, D. A. A. Pertiwi, and Y. Dasril, “New model combination meta-learner to improve accuracy prediction P2P lending with stacking ensemble learning,” *Intell. Syst. with Appl.*, vol. 18, p. 200204, 2023.
[23] D. I. Wijaya, M. K. Aulia, J. Jumanto, and M. F. Al Hakim, “Room occupancy classification using multilayer perceptron,” *J. Soft Comput. Explor.*, vol. 2, no. 2, pp. 163–168, 2021.
[24] V. Chang, V. R. Bhavani, A. Q. Xu, and M. A. Hossain, “An artificial intelligence model for heart disease detection using machine learning algorithms,” *Healthc. Anal.*, vol. 2, p. 100016, 2022.
[25] A. A. Almazroi, E. A. Aldhahri, S. Bashir, and S. Ashfaq, “A Clinical Decision Support System for Heart Disease Prediction using Deep Learning,” *IEEE Access*, 2023.
[26] A. A. Hussein, “Improve the performance of K-means by using genetic algorithm for classification heart attack,” *Int. J. Electr. Comput. Eng.*, vol. 8, no. 2, p. 1256, 2018.
[27] M. Ozcan and S. Peker, “A classification and regression tree algorithm for heart disease modeling and prediction,” *Healthc. Anal.*, vol. 3, p. 100130, 2023.
[28] M. Gjoreski, A. Gradišek, B. Budna, M. Gams, and G. Poglajen, “Machine learning and end-to-end deep learning for the detection of chronic heart failure from heart sounds,” *Ieee Access*, vol. 8, pp. 20313–20324, 2020.
[29] S. A. Ali *et al.*, “An optimally configured and improved deep belief network (OCI-DBN) approach for heart disease prediction based on Ruzzo–Tompa and stacked genetic algorithm,” *IEEE Access*, vol. 8, pp. 65947–65958, 2020.