Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

Comparison of Machine Learning Approaches for the Modelling of Non-Invasive Optical-Based Hemoglobin Measurement Device

Karlisa Priandana[1,8,*], Afif Mucholis[1], Rinaldy Permana Lubis[1], Irzaman[2], Sri Wahjuni[1], Renan Prasta Jenie[3], Naufal Muharam Nurdin[4,7], M Dahrul[2,5], Arga Aridarma[6], Husin Alatas[2,8]

[1] Computer Science Study Program, School of Data Science, Mathematics and Informatics, IPB University, Bogor 16680, Indonesia
[2] Department of Physics, Faculty of Mathematics and Natural Sciences, IPB University, Bogor 16680, Indonesia
[3] Department of Master of Public Health, Binawan University, Jakarta 13630, Indonesia
[4] Department of Community Nutrition, IPB University, Bogor 16680, Indonesia
[5] Department of Electronic and Computer Engineering, College of Electrical Engineering and Computer Science, National Taiwan University of Science and Technology, Taipei City 106335, Taiwan (R.O.C.)
[6] PT Tesena Inovindo, Jakarta, 13750, Indonesia
[7] Faculty of Medicine, IPB University, Bogor 16680, Indonesia
[8] Center for Transdisciplinary and Sustainability Sciences (CTSS). IPB University, Bogor 16127, Indonesia

ABSTRACT

This paper discusses the comparison of several machine learning approaches to build a model for our developed non-invasive optical-based hemoglobin measurement device. The optical data used for model development was obtained from Rumah Sakit Universitas Brawijaya (RSUB) Malang. The dataset includes information from patients aged 18 and above who underwent hemoglobin testing at RSUB between August and October 2022. A total of 610 patients consented to participate in the study by signing a consent form. The ground truth Hb values, used as data labels, were obtained from blood tests conducted in the RSUB laboratory. The features used for predicting Hb values were derived from the light reflections captured by a non-invasive hemoglobin measurement device that was developed for this study. Modeling was performed using two machine learning methods: Multi-Variable Linear Regression (MVLR) and Random Forest Regressor. The MVLR model resulted in five important features with a Mean Absolute Error (MAE) of 1.37 g/dL, while the Random Forest Regressor model, with 20 features and a tree depth of 8, achieved an MAE of 0.97 g/dL. Based on these results, the Random Forest Regressor model is recommended for implementation in the device. This model can be considered to meet industry standards, as the MAE value is below 1.5 g/dL.

Keywords:

Hemoglobin; multi-variable linear regression; random forest; regression; univariate feature selection

1. Introduction

Measurement of blood hemoglobin level is important to diagnose if someone has anemia. According to the World Health Organization (WHO), someone is diagnosed with anemia if the hemoglobin (Hb) level is below 12.0 g/dL for women or below 13.0 g/dL for men [1]. It is also found

* *Corresponding author.*
E-mail address: karlisa@apps.ipb.ac.id

https://doi.org/10.37934/araset.63.1.5266

that the hemoglobin level in pregnancy significantly determine the health of mother and fetus, where severe anemia may associate with prematurity, spontaneous abortions, low birth weight, and fetal deaths [2]. Due to this issue, the measurement of hemoglobin level for pregnant women has been one of the main concerns in various countries, including in Indonesia [3], Malaysia [4], and Japan [5]. There are some methods that can be used to measure hemoglobin level. However, most of the methods are invasive methods, *i.e.*, blood sample is taken from the patient and hemoglobin level is calculated from the blood sample in a laboratory [6]. Although invasive methods can measure the hemoglobin level accurately, some people may hesitate to measure their hemoglobin level since the procedure not only causes pain but also increases the possibility for someone to be infected by diseases.

Recently, advancements in non-invasive healthcare monitoring devices have led to the development of technologies for tracking blood flow circulation [7], vital body signals [8], and predicting kidney diseases [9]. Additionally, devices for non-invasive hemoglobin measurement have also emerged [10, 11]. These devices measure hemoglobin levels in real-time using photometric pulse technology [12]. In principle, the photo-spectroscopy works by the excitation of light on a substance. Then, the reflected light is used to indicate the response of the examined material, *i.e.*, the patients' finger. The intensity of the reflected light received by the sensor is then used to determine the hemoglobin level of the patient. For this purpose, an algorithm is required in order to determine the hemoglobin level based on the reflected light intensity. Some literatures have also compared the performance of invasive and noninvasive blood hemoglobin measurements [13-15]. It was noted that the noninvasive hemoglobin measurement provides acceptable accuracy compared to that of the standard invasive method.

The main challenge of non-invasive blood hemoglobin measurement device development is to find the function between the reflected light intensity (input) and hemoglobin level (output). The most straightforward approach to answer this issue is by developing a model using machine learning method, such as multi-model stacking regressor [16], deep learning [17] and other regression-based machine learning approaches [18]. We have developed a portable and lightweight non-invasive blood hemoglobin measurement device, where we employed a linear regression model, as reported in our previous work [10]. However, the previous work built the model based on very few test subjects and low data variability. Thus, some data properties may not be captured well and the model may be over-simplified.

In this paper, we conducted an experiment to acquire data from 610 volunteering patients who had their blood drawn at Rumah Sakit Universitas Brawijaya (RSUB) Malang between August and October 2022 to check their blood Hemoglobin levels. All patients agreed to have their data collected using the developed non-invasive Hb measurement device, as documented in consent forms. In the experiment, each patient's optical data was recorded five times. Then, machine-learning models were generated by using the reflected optical data from the patients' fingers and their corresponding lab-test blood hemoglobin level. In this study, we explore the possibility of using more advanced machine learning algorithms [19] in order to determine the best model for our non-invasive hemoglobin measurement device. Two algorithms were analyzed and compared, namely Multi-Variable Linear Regression (MVLR) and Random Forest (RF).

2. Methodology

2.1 Non-Invasive Device for Data Acquisition

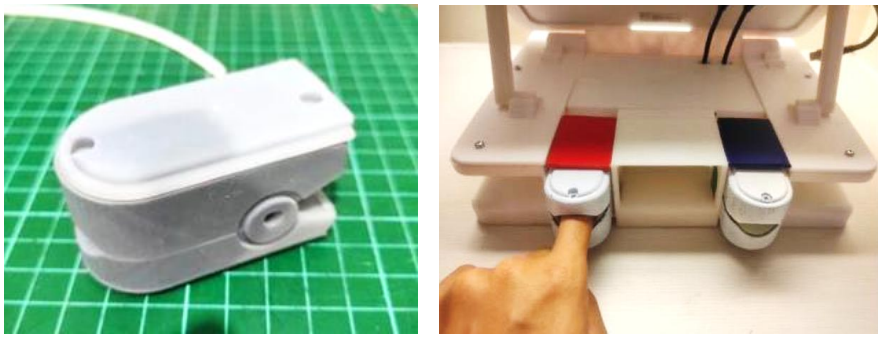
Data acquisition was conducted using our developed non-invasive blood hemoglobin level measurement device, as shown in Figure 1(a). The device is lightweight and portable, allowing

measurements to be taken anytime, similar to other lightweight non-invasive medical devices [7, 10]. Our device utilizes two LEDs, each emitting light at wavelengths of 680 nm and 810 nm. These wavelengths differ from those used in previous studies (660 nm and 940 nm) [20], based on results from further laboratory experiments. Optical data for hemoglobin (Hb) prediction was obtained from 610 volunteering patients, with each patient's finger placed inside the device (see Figure 1(b)). The device used a sensor to capture the reflected signal intensity in four different settings, as illustrated in Figure 2:

i. Both LEDs off
ii. LED 1 on
iii. LED 2 on
iv. Both LEDs on

Each setting was further varied by adjusting the LED on/off duration, resulting in 10 different variations for each setting. Data were collected 200 times for each variation. Thus, the total number of optical signal features obtained was 4 settings × 10 durations × 200 measurements = 8000 optical signal features.

The reflected light intensity was measured 5 times per patient, resulting in a total of 3050 optical data points (610 patients × 5 measurements) in the dataset. The gold standard for these data, used for model training, was obtained from invasive laboratory hemoglobin (Hb) tests conducted at RSUB Malang on the same day as the optical data measurement.



(a) (b)

Fig. 1. Non-invasive blood hemoglobin level measurement device (a) The developed prototype (b) Optical data measurement

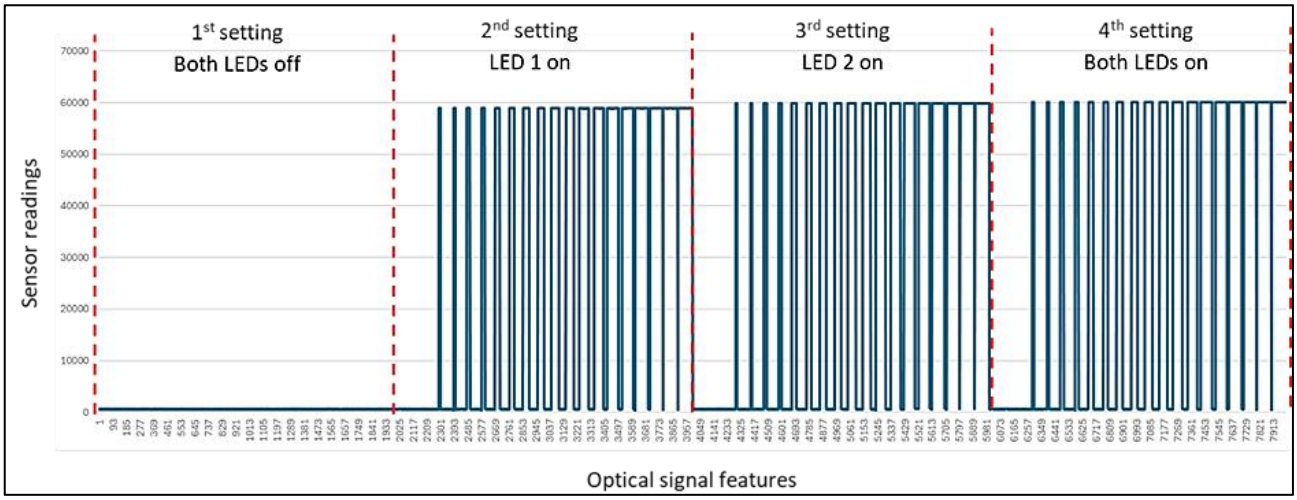


Fig. 2. Sample of reflected signal intensity captured by the non-invasive blood hemoglobin measurement device

2.2 Data Description

The data used in this study were the hemoglobin (Hb) measurement data which was carried out on 610 patients, collected from August 2022 to October 2022. The ground truth of blood Hb levels as the data label was obtained from blood tests in the laboratory. The data features/attributes for classifying the blood Hb levels are the reflected optical light from the probe obtained from the patients' finger using the non-invasive hemoglobin measurement device that has been developed (Figure 1(b)). In this experiment, data acquisition was carried out using four devices with the following details:

i. Device 1 = 965 data points (193 patients)
ii. Device 2 = 1540 data points (308 patients)
iii. Device 4 = 190 data points (38 patients)
iv. Device 5 = 355 data points (71 patients)

The generated dataset was transformed into a tabular format, comprising a total of 3050 data points (rows), which were collected from 610 respondents with 5 repetitions each. The dataset includes 8000 attributes (columns), representing the features extracted from the optical signal measurements.

Before processing the data, the correlation between the features/attributes of the optical data and the data labels (blood Hb level) is calculated using Pearson Correlation method. The maximum absolute correlation result for the data is 0.2996, as shown in Figure 3. The number of data points for each Hb value can be seen in Figure 4. It can be observed that the Hb values measured at RSUB from August to October 2022 range from 6 to 19.9 g/dL, demonstrating considerable variability.

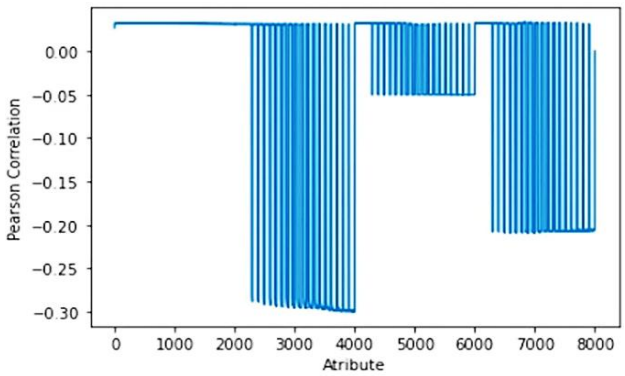


Fig. 3. Correlation of data attributes with Hb values (maximum absolute correlation value = 0.2996)

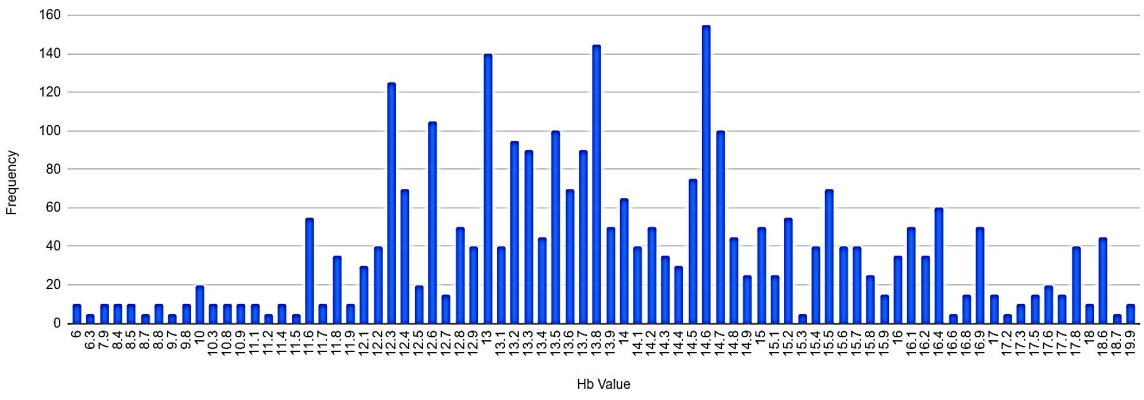


Fig. 4. Data frequency per hemoglobin value

2.3 Modelling of Non-Invasive Optical-Based Hemoglobin Measurement Device

Data processing was conducted for the obtained data (3050 samples) as shown in Figure 5. The dataset was divided into training and test datasets with an 80:20 ratio, resulting in a total of 2440 training samples and 610 test samples. The training data was then used to train the models for predicting Hb values based on the data attributes/features. Based on previous analyses, the algorithms used for modeling are as follows:

i. Multi-Variable Linear Regression
ii. Random Forest Regressor

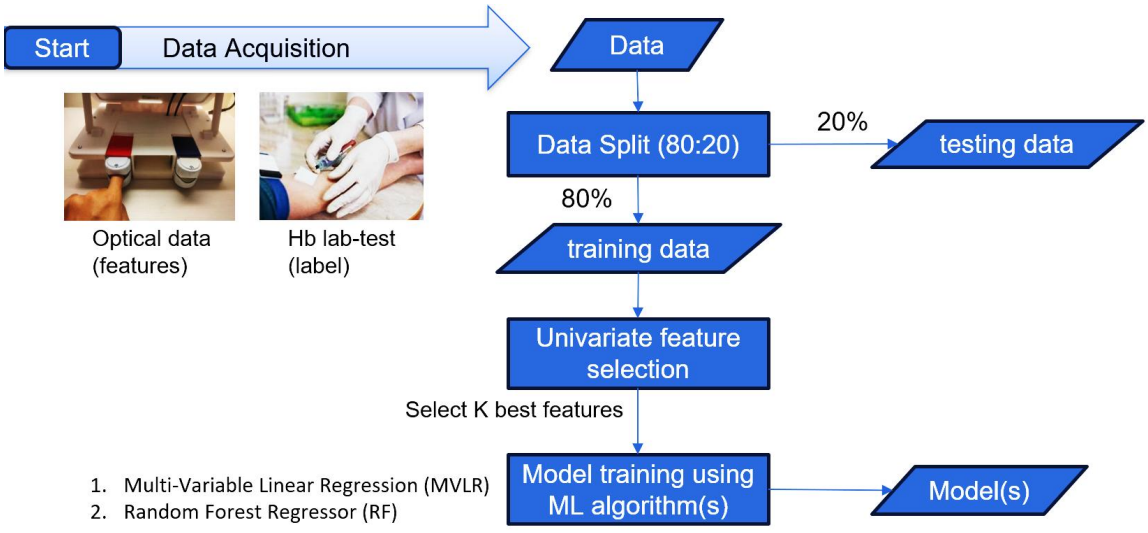


Fig. 5. Data processing

2.3.1 Univariate feature selection

All modeling was performed by selecting several important features using univariate feature selection. This method works by choosing the best features based on univariate statistical tests [21]. By focusing on the significance of individual features, univariate feature selection helps in filtering out irrelevant or redundant features, which can lead to more efficient and effective models [22]. The process of univariate feature selection involves applying statistical tests to measure the relationship between each feature and the target variable. Commonly used tests include the Chi-Square test for categorical features, ANOVA (Analysis of Variance) for continuous features, and the correlation coefficient for assessing linear relationships [23, 24].

After identifying the best features, this study utilized SelectKBest method for selecting K most important features and removing all other features. In this study, various values of K were tested, ranging from 5 to 100 features in increments of 5.

2.3.2 Multi-variable linear regression

Multivariable linear regression algorithm produces a model that is used to determine the function / the relationship between 1 dependent variable, *i.e.* the data label, and more than 1 independent variable. Multivariable linear regression is one of the most widely used algorithms among the other algorithms such as the multivariable logistic regression and Cox proportional

hazards regression [25]. Multivariable linear regression works by taking the linear relationship between the data label and some data attributes by using the following equation:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_K X_K + \varepsilon \quad (1)$$

where Y is the dependent variable, and $X_1\ X_2\ X_3 \ldots X_K$ are K independent variables (as selected by the feature selection method, for example). When K = 1, equation (1) becomes a linear regression model. In the equation, ε is the random error, β_0 is the regression constant, and $\beta_1\ \beta_2 \ldots \beta_K$ are the unknown values to be determined by the algorithm.

2.3.3 Random forest regressor

Random Forest Regressor is a machine learning model used to predict continuous values based on multiple input features. It operates by constructing an ensemble of decision trees during the training phase. This ensemble approach enhances the model's prediction accuracy and robustness, as outlined by [26]. To build these trees, the model employs bootstrap sampling. This technique involves generating multiple decision trees from bootstrap samples of the original training data. Each bootstrap sample is created by randomly drawing N samples with replacement from the original dataset, producing B bootstrap samples in total $\{S_1, S_2, \ldots, S_B\}$. For each bootstrap sample S_b where b = 1, 2, …, B, a decision tree T_b is built [27]. At each node of the tree, a random subset of m features is randomly chosen from the total available M features. This randomness helps ensure that the trees are diverse and reduces the chance of overfitting, as explained by Liaw and Wiener (2002) [28].

The rule of thumb for this subset value in regression is $m = M/3$. Each decision tree is trained to predict the target value based on the input features. In the case of regression, the trees are optimized to minimize prediction errors using metrics such as mean squared error (MSE) or mean absolute error (MAE) [26]. Once the decision trees are built, the Random Forest Regressor makes predictions by averaging the predictions from all the trees, as defined in the following equation:

$$\hat{y} = \frac{1}{B}\sum_{b=0}^{B} T_b(x) \quad (2)$$

As in Eq. (2), the final prediction for new data from the Random Forest Regressor is obtained by averaging individual trees' predictions, leveraging the diversity among the trees to enhance performance [28]. This aggregation process also helps to smooth out individual trees' errors and improves the overall prediction accuracy.

3. Results

3.1 Prediction of Blood Hb Level Using Multi-Variable Linear Regression (MVLR)

Modeling using Multi-Variable Linear Regression was performed with the training data. Once the model was obtained, it was used to predict Hb values from the test data. The predicted Hb values from the test data generated by the model were compared with the actual Hb values, and the error was evaluated using Mean Absolute Error (MAE). This experiment included modeling for each individual device used in data collection (devices 1, 2, 3, and 4) as well as for the combined data from all devices. MAE calculations were performed for both the training and test data. Additionally, MAE was calculated separately for data from each device (1, 2, 3, and 4) and for the

combined data from all devices. The MAE values relative to the number of features can be seen in Figures 6(a) to 6(d).

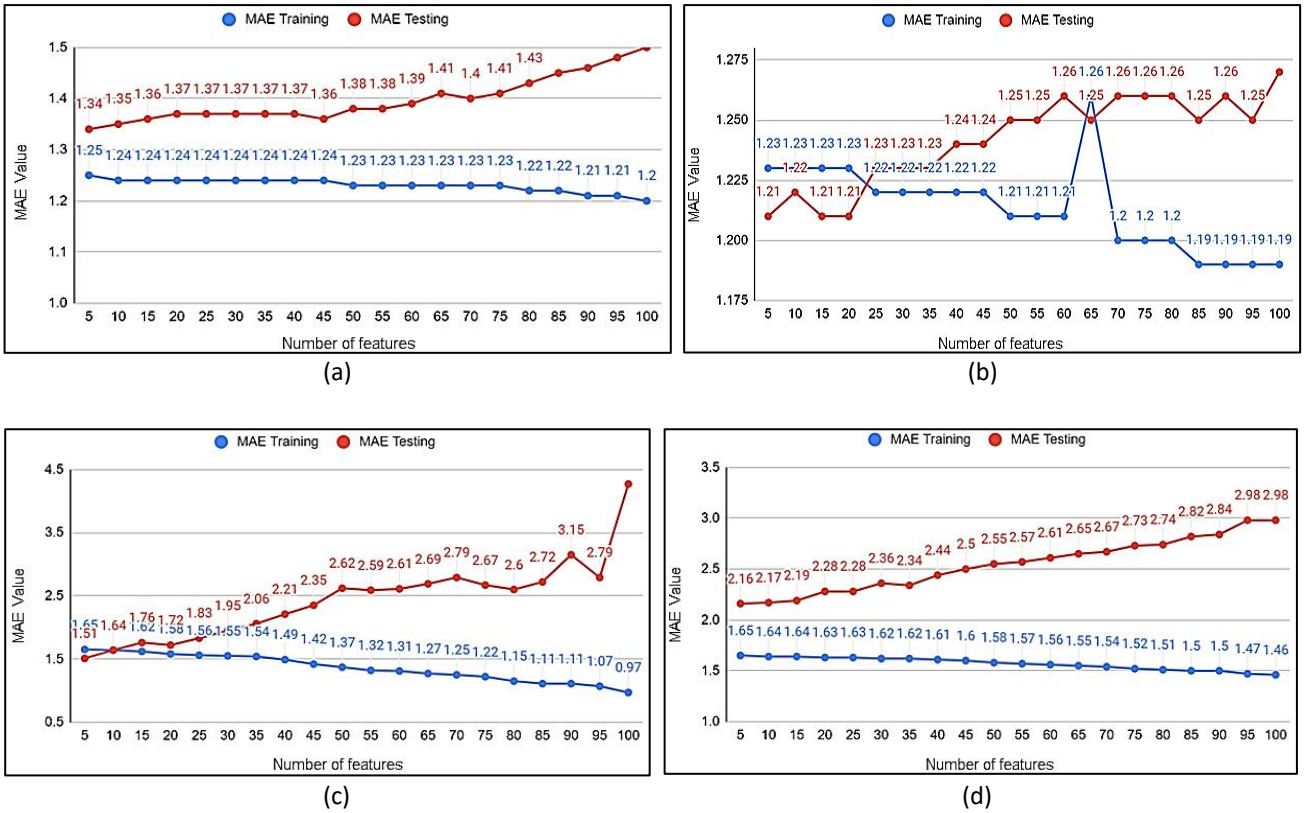
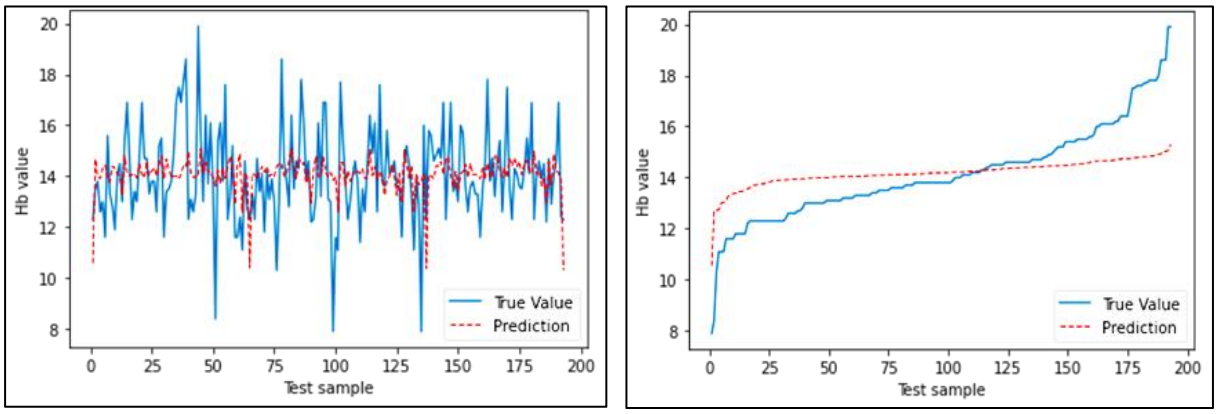


Fig. 6. Trend of MAE values with respect to the number of features = 5 (a) Devices 1 (lowest testing MAE = 1.34 g/dL) (b) Devices 2 (lowest testing MAE = 1.21 g/dL) (c) Device 3 (lowest testing MAE = 1.51 g/dL) (d) (lowest testing MAE = 2.16 g/dL)

Based on Figure 6, it can be observed that as the number of features increases, the MAE for the training data decreases, while the MAE for the test data increases. This indicates that the model is overfitting to the training data. Therefore, a model is chosen where the MAE is optimized for both training and test data. This optimal value is achieved with 5 features. Choosing a small number of features not only helps avoid overfitting but also facilitates implementation in the device. Consequently, the top 5 features are selected, with each device showing the MAE for training and testing as indicated in Table 1. Based on Table 1, device 2 has the smallest MAE for both training and testing compared to the other three devices. Furthermore, the Hb prediction results using Multi-Variable Linear Regression for each device can be seen in Figures 7 to 10.

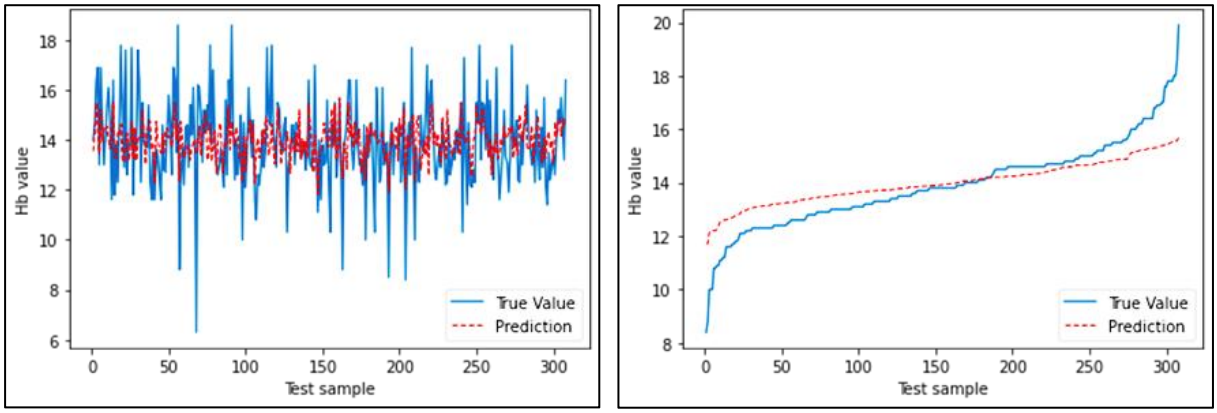
Table 1
Training and testing MAE values (in g/dL) for each device using the top 5 features

Device	MAE train	MAE test
1	1.25	1.31
2	1.23	1.21
3	1.65	1.51
4	1.65	2.16



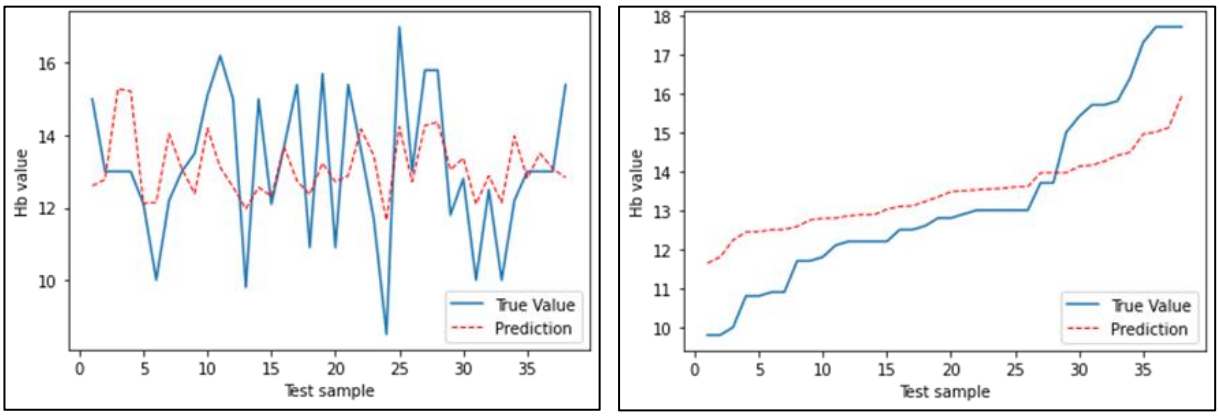
(a) (b)

Fig. 7. Hb level prediction results using MVLR for device 1 (MAE = 1.31. g/dL) (a) Unsorted (b) Sorted



(a) (b)

Fig. 8. Hb level prediction results using MVLR for device 2 (MAE = 1.21 g/dL) (a) Unsorted (b) Sorted

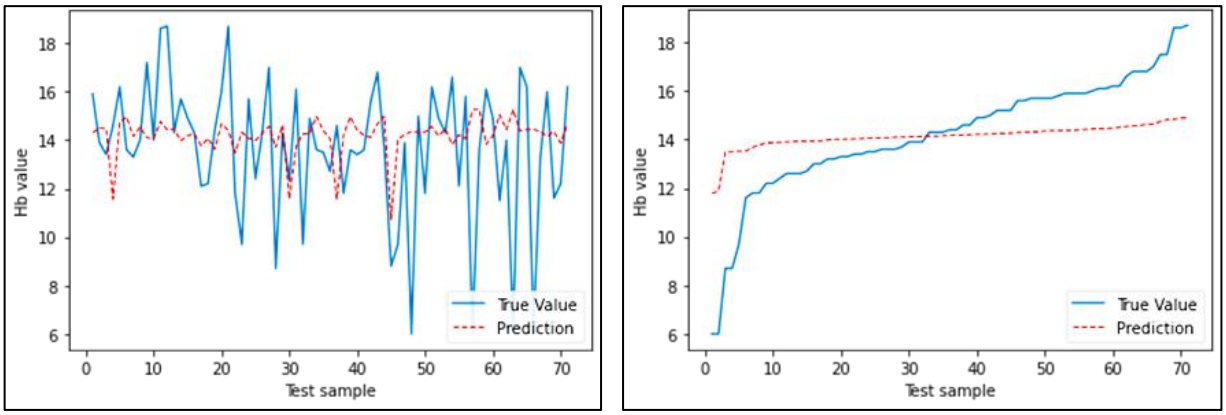


(a) (b)

Fig. 9. Hb level prediction results using MVLR for device 3 (MAE = 1.51 g/dL) (a) Unsorted (b) Sorted

From the modeling results of each device, it can be concluded that all devices can be modeled quite effectively. Therefore, in addition to modeling each device individually, data from all devices was also used to create a general model. This is necessary because, ultimately, the general model will be implemented in the developed non-invasive Hb measurement device. As before, an

experiment was conducted to determine the most important features. The trend of MAE values with respect to the number of important features used can be seen in Figure 11.



(a) (b)

Fig. 10. Hb level prediction results using MVLR for device 4 (MAE = 2.16 g/dL) (a) Unsorted (b) Sorted

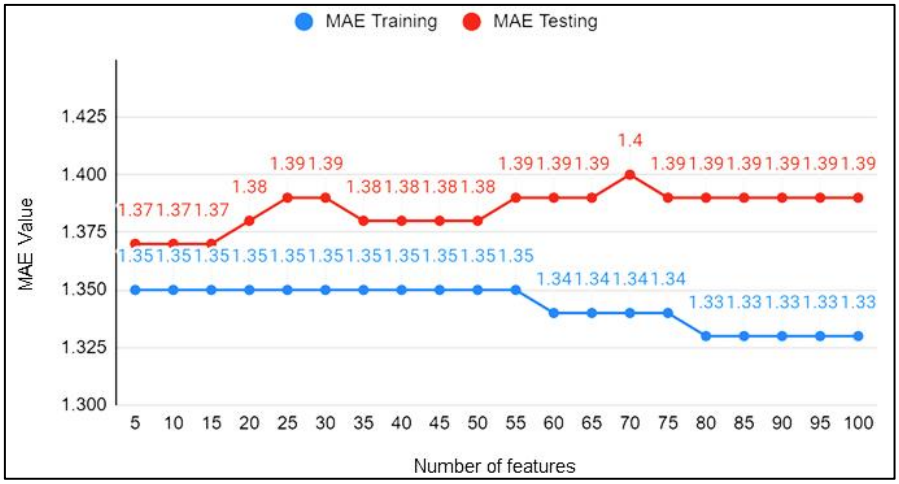
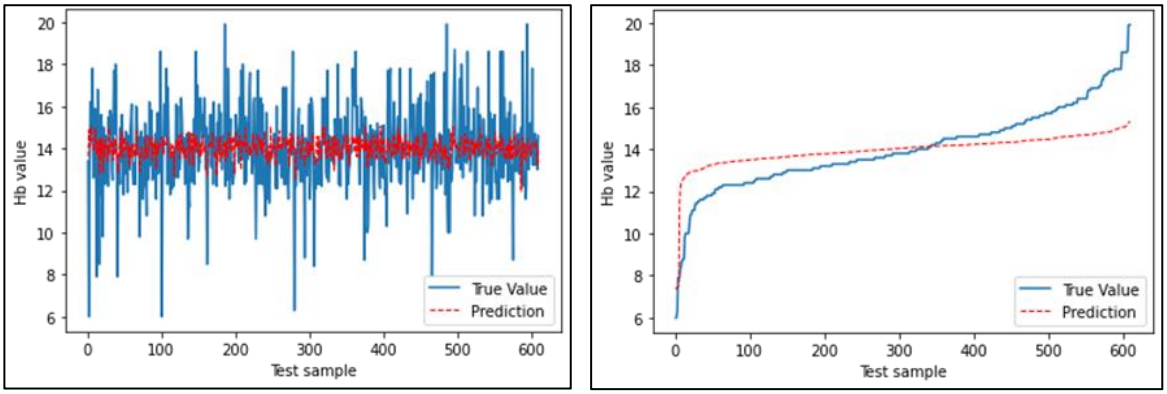


Fig. 11. Trend of MAE values with respect to the number of features for all data (lowest testing MAE = 1.37 g/dL; number of features = 5)

Based on Figure 11, 5 features were selected. In this case, the MAE values obtained are a training MAE of 1.35 g/dL and a testing MAE of 1.37 g/dL. The coefficient values obtained for these 5 features are shown in Table 2. The Hb prediction results for the 5 most important features can be seen in Figure 12.

Table 2
Coefficients of multi-variable linear regression for the most important features

i	Feature number (x_i)	Coefficient (a_i)
0	-	115.1623518
1	3994	-0.01160635
2	3899	-0.00459493
3	3992	0.00300189
4	3990	0.01186458
5	3893	-0.00038707



(a) (b)

Fig. 12. Hb level prediction results using MVLR (MAE = 1.37 g/dL) (a) Unsorted (b) Sorted

3.2 Prediction of Hb Level Using Random Forest Regressor (RF)

Random Forest is a learning method that operates by building multiple decision trees during the training phase. In this phase, modeling was performed with varying tree depths, specifically at values of 2, 4, 6, 8, and 10. The resulting models were then used to predict Hb values from the test data. The predicted Hb values from the test data generated by the models were compared with the actual Hb values, and the errors were evaluated using Mean Absolute Error (MAE). The trends of training MAE and testing MAE with respect to the number of features and tree depth for all devices can be seen in Figures 13 and 14.

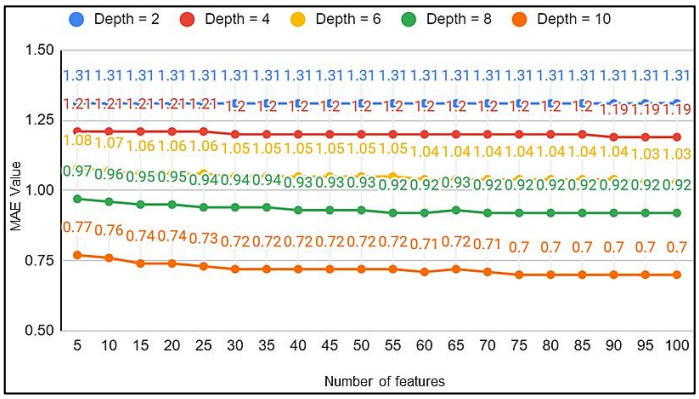


Fig. 13. Trend of training MAE with respect to the number of features and tree depth using Random Forest Regressor

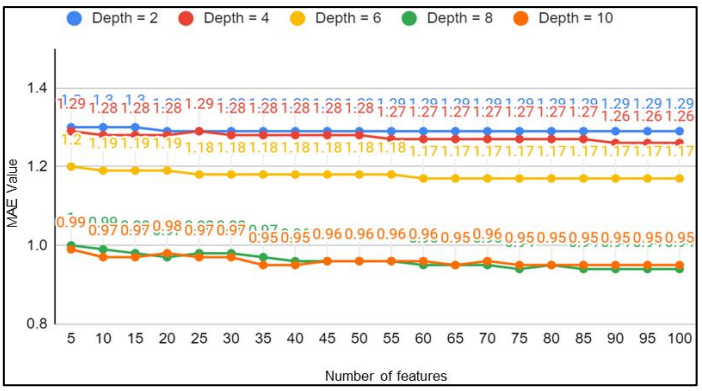


Fig. 14. Trend of testing MAE with respect to the number of features and tree depth using Random Forest Regressor

Based on Figure 14, it can be observed that tree depths of 8 and 10 yield overlapping MAE values. Therefore, a depth of 8 was chosen. Details of the MAE values for depths of 8 and 10 can be seen in Table 3. Based on Table 3, a relatively low MAE value was obtained with 20 features and a depth of 8, resulting in a test MAE of 0.97 g/dL. The 20 features selected by the algorithm are listed in Table 4. The Hb prediction results using the Random Forest Regressor method with 20 features and a depth of 8 can be seen in Figure 15.

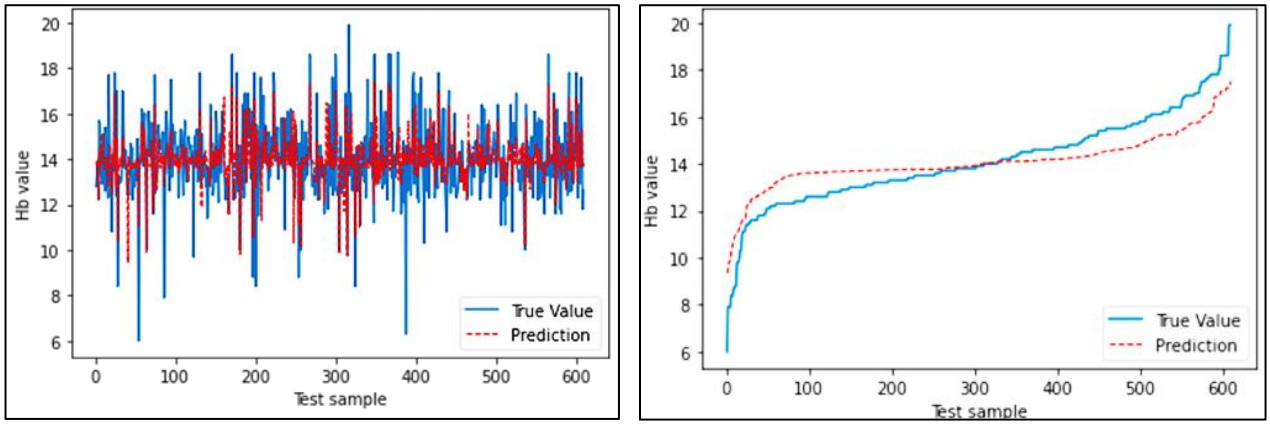
Table 3
Testing MAE values using the Random Forest Regressor model (depth = 8 and depth = 10)

Number of Features	Nilai MAE testing (in g/dL)	
	Depth = 8	Depth = 10
5	1	0.99
10	0.99	0.97
15	0.98	0.97
20	0.97	0.98
25	0.98	0.97
30	0.98	0.97
35	0.97	0.95
40	0.96	0.95
45	0.96	0.96
50	0.96	0.96
55	0.96	0.96
60	0.95	0.96
65	0.95	0.95
70	0.95	0.96
75	0.94	0.95
80	0.95	0.95
85	0.94	0.95
90	0.94	0.95
95	0.94	0.95
100	0.94	0.95

Table 4
The 20 most important featured used in the Random Forest Regressor modelling

Rank	Feature no.	Rank	Feature no.	Rank	Feature no.	Rank	Feature no.
1	3994	6	4000	11	3892	16	3988
2	3899	7	3987	12	3970	17	3852
3	3992	8	3939	13	3799	18	3876
4	3990	9	3963	14	3979	19	3841
5	3893	10	3790	15	3934	20	3984

Next, predictions using the Random Forest Regressor model with 20 features and a depth of 8 were also made for each repetition. In this case, each patient's data was collected 5 times. Each repetition was then predicted using the trained model. The prediction results for the five repetitions can be seen in Table 5 and Figure 16.



(a) (b)

Fig. 15. Hb level prediction results using Random Forest Regressor with number of features = 20 and depth = 8 (MAE = 0.97 g/dL) (a) Unsorted (b) Sorted

Table 5
Hb prediction results using Random Forest Regressor for the five repetitions of data measurement

No	Hb gold standard (g/dL)	Measurement repetition no. (Hb prediction in g/dL)					Precision value	Average measurement (g/dL)	Measurement deviation (g/dL)
		1	2	3	4	5			
1	6	9.79	8.73	9.26	9.26	8.61	5.19%	9.13	3.13
2	6.3	9.99	9.78	9.62	9.62	9.96	1.85%	9.79	3.49
3	7.9	10.00	9.79	9.78	9.78	9.98	1.14%	9.87	1.97
4	8.4	11.01	10.00	9.96	9.96	9.99	4.51%	10.18	1.78
5	8.8	11.10	10.10	10.02	10.02	10.06	4.58%	10.26	1.46
6	9.7	11.49	10.26	10.08	10.08	10.08	5.89%	10.40	0.70
7	9.8	11.70	10.28	10.10	10.10	11.35	7.13%	10.71	0.91
8	10	12.26	11.62	11.80	11.80	11.85	1.99%	11.87	1.87
9	10.8	12.84	11.87	11.97	11.97	12.06	3.27%	12.14	1.34
10	11.1	13.04	12.11	11.97	11.97	12.22	3.65%	12.26	1.16
11	11.5	13.14	12.36	12.26	12.26	12.26	3.10%	12.46	0.96
12	12.2	13.51	13.51	13.51	13.51	13.51	0.01%	13.51	1.31
13	12.7	13.62	13.62	13.60	13.60	13.70	0.28%	13.63	0.93
14	13.1	13.71	13.71	13.68	13.68	13.70	0.10%	13.69	0.59
15	13.5	13.77	13.77	13.76	13.76	13.85	0.27%	13.78	0.28
16	14.1	14.12	14.12	14.14	14.14	14.13	0.06%	14.13	0.03
17	14.7	14.12	14.13	14.15	14.15	14.13	0.08%	14.14	0.56
18	15.1	14.64	14.62	14.65	14.65	14.67	0.12%	14.65	0.45
19	15.6	14.66	14.66	14.68	14.68	14.69	0.10%	14.67	0.93
20	16.1	14.75	14.77	14.76	14.76	14.77	0.04%	14.76	1.34
21	16.8	15.10	15.04	15.19	15.19	15.28	0.62%	15.16	1.64
22	17.2	15.17	15.05	15.35	15.35	15.48	1.11%	15.28	1.92
23	17.7	15.18	15.11	15.61	15.61	15.69	1.76%	15.44	2.26
24	18	15.30	15.62	15.72	15.72	16.49	2.76%	15.77	2.23
25	18.7	15.51	15.70	15.83	15.83	16.82	3.21%	15.94	2.76
26	19.9	15.93	16.05	16.38	16.38	18.40	6.09%	16.63	3.27
Average							2.27%	MAE = 1.51 g/dL; R2 = 0.789	
Average for measurements within 8.8 g/dL – 16.8 g/dL (rows 5 to 21)							1.84%	MAE = 0.97 g/dL; R2 = 0.797	

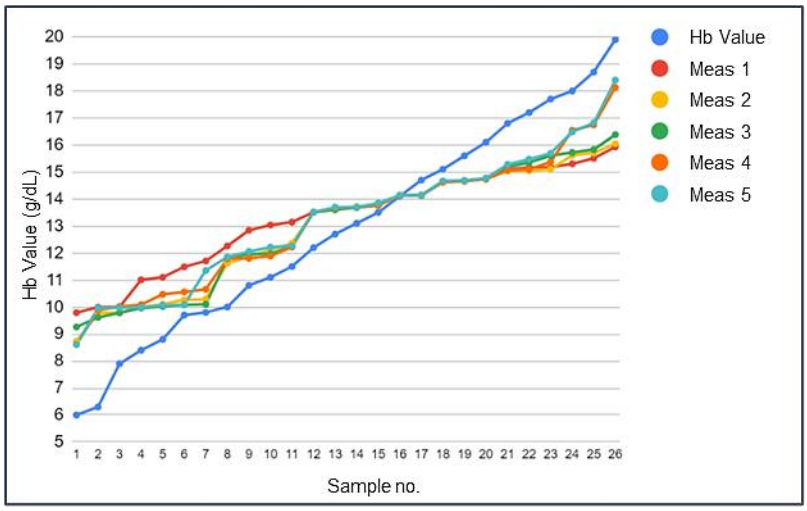


Fig. 16. Hb level prediction results using Random Forest Regressor for five repetitions (average deviation = 2.27%)

Based on Table 5 and Figure 16, the average measurement deviation (absolute difference between predicted values and true values) is 1.51 g/dL, while the average precision of the device for the 5 repetitions is 2.27%. The corresponding R2 value for this performance is 0.789. The performance characteristics of the target device are shown in Table 6. According to these standards, the measurement range that meets the target is 8.8 g/dL – 16.8 g/dL (highlighted in yellow and bold in Table 5). Using this measurement range, the average measurement deviation by the device (absolute difference between predicted values and true values) is 0.97 g/dL, and the device precision is 1.84%. The corresponding R2 value for this performance is 0.797. Thus, this model is recommended to be implemented in the device.

Table 6
Targeted performance of the developed non-invasive blood Hemoglobin measurement device

Performance characteristic	Optimum	Minimum
Accuracy	±1 g/dL	±1.75 g/dL
Precision	1.5%	2%

4. Conclusions

This study analyzed and compared machine learning algorithms to build the best model for a non-invasive blood hemoglobin measurement device. Two algorithms were analyzed and compared, namely the Multi-Variable Linear Regression (MVLR) and Random Forest Regressor (RF). For this purpose, we conducted an experiment to acquire data from 610 volunteering patients who had their blood drawn at Rumah Sakit Universitas Brawijaya (RSUB) Malang between August and October 2022 to check their blood hemoglobin levels. In the experiment, each patient's optical data was recorded. Then, machine-learning models were generated by using the reflected optical data from the patients' fingers and their corresponding lab-test blood hemoglobin level.

The data modeling with Multi-Variable Linear Regression (MVLR) identified 5 key features, resulting in a Mean Absolute Error (MAE) of 1.37 g/dL. In comparison, the Random Forest Regressor, utilizing 20 features and a tree depth of 8, achieved a lower MAE of 0.97 g/dL. The corresponding R2 value for this performance is 0.797. Therefore, the Random Forest Regressor model is

recommended for implementation in the device, as it meets industry standards with an MAE below 1.5 g/dL.

Further analysis of the Random Forest Regressor model, which was evaluated through multiple measurement repetitions, revealed an average measurement deviation of 1.51 g/dL (absolute difference between predicted and true values) and a device precision of 2.27%. According to the performance standards, the measurement range that better meets the targeted performance is between 8.8 g/dL and 16.8 g/dL. Within this range, the average measurement deviation is reduced to 0.97 g/dL, and the device precision improves to 1.84%.

This study is part of the development of non-invasive blood hemoglobin measurement device. The developed device will be implemented as a ready-to-use device through multidisciplinary research collaboration. Future research tasks include: (i) embedding the model to the device, (ii) testing the device and (iii) commercialization.

Acknowledgement

The authors would like to express their gratitude to the Education Fund Management Institution (Lembaga Pengelola Dana Pendidikan, LPDP) - Ministry of Finance of the Republic of Indonesia, for the Rispro LPDP third-year research grant No. S-325/LPDP.4/2023, titled “Kemandirian Bangsa Dalam Teknologi Proses Pembuatan Alat Pengukuran Kadar Hb Darah Non Invasif.” We also extend our thanks to the Directorate of Research and Innovation, IPB University, for their support in conducting this research. Additionally, we appreciate Rumah Sakit Universitas Brawijaya (RSUB) Malang for their assistance in collecting the dataset.

References

[1] Cappellini, M. Domenica, and Irene Motta. "Anemia in clinical practice—definition and classification: Does hemoglobin change with aging?." In *Seminars in hematology*, vol. 52, no. 4, pp. 261-269. WB Saunders, 2015. https://doi.org/10.1053/j.seminhematol.2015.07.006

[2] Sifakis, S., and G. Pharmakides. "Anemia in pregnancy." *Annals of the New York Academy of Sciences* 900, no. 1 (2000): 125-136. https://doi.org/10.1111/j.1749-6632.2000.tb06223.x

[3] Darmawati, Darmawati, Tongku Nizwan-Siregar, Kamil Hajjul, and Tahlil Teuku. "Exploring Indonesian mothers' perspectives on anemia during pregnancy: A qualitative approach." *Enfermería Clínica (English Edition)* 32 (2022): S31-S37. https://doi.org/10.1016/j.enfcle.2020.11.007

[4] Abd Rahman, Raudah, Idayu Badilla Idris, Zaleha Md Isa, Rahana Abdul Rahman, and Zaleha Abdullah Mahdy. "The prevalence and risk factors of iron deficiency anemia among pregnant women in Malaysia: A systematic review." *Frontiers in nutrition* 9 (2022): 847693. https://doi.org/10.3389/fnut.2022.847693

[5] Noshiro, Kiwamu, Takeshi Umazume, Rifumi Hattori, Soromon Kataoka, Takashi Yamada, and Hidemichi Watari. "Hemoglobin concentration during early pregnancy as an accurate predictor of anemia during late pregnancy." *Nutrients* 14, no. 4 (2022): 839. https://doi.org/10.3390/nu14040839

[6] Whitehead Jr, Ralph D., Zuguo Mei, Carine Mapango, and Maria Elena D. Jefferds. "Methods and analyzers for hemoglobin measurement in clinical laboratories and field settings." *Annals of the New York Academy of Sciences* 1450, no. 1 (2019): 147-171. https://doi.org/10.1111/nyas.14124

[7] Adib, Mohd Azrul Hisham Mohd, Khairul Shah Affendy Yakof, Zulfikar Anuar, Nizam Abdullah, Idris Mat Sahat, and Nur Hazreen Mohd Hasni. "The cardio vascular simulator (CardioVASS) device on monitoring the physiology of blood flow circulation via angiographic image for medical trainee." *Journal of Advanced Research in Fluid Mechanics and Thermal Sciences* 62, no. 1 (2019): 103-111.

[8] Madona, Putri, and Jepri Simatupang. "Multisensory health monitoring device based on raspberry Pi 4B." *Journal of Advanced Research in Applied Sciences and Engineering Technology* 42, no. 1 (2024): 42-56. https://doi.org/10.37934/araset.42.1.4256

[9] Gopi, Rahul Sanmugam, Tavanam Venkata Rao, Uma Maheswari Sethuramasamy, Jeyasekaran Jeno Jasmine, U. Soma Naidu, and Sundararajan Edwin Raja. "An efficient IoT-artificial intelligence-based chronic kidney diseases prediction using temporal convolutional network (TCN) deep learning method in healthcare system." *Journal of Advanced Research in Applied Sciences and Engineering Technology* 32, no. 3 (2023): 165-177. https://doi.org/10.37934/araset.32.3.165177

[10] Hardyanto, Ichsan, Sabar Pambudi, Yaya Suyarna, Arga Ardidarma, Ade Kurniawan, Johan Iskandar, Ridwan Siskandar, Renan Prasta Jenie, Husin Alatas, and Irzaman Irzaman. "Non-invasive hemoglobin blood level measurement system." In *AIP Conference Proceedings* 2320, no. 1. AIP Publishing, 2021. https://doi.org/10.1063/5.0037659

[11] Rochmanto, Raditya Artha, Hasballah Zakaria, Ratih Devi Alviana, and Nurhalim Shahib. "Non-invasive hemoglobin measurement for anemia diagnosis." In *2017 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, pp. 1-5. IEEE, 2017. https://doi.org/10.1109/EECSI.2017.8239096

[12] Irzaman, Irzaman, Yaya Suryana, Sabar Pambudi, Tika Widayanti, Renan Prasta Jenie, Bayu Prastowo, Refky Zaheri, Ichsan Hardyanto, Naufal Muharam Nurdin, Muhammad Dahrul, Johan Iskandar, Ade Kurniawan, Ridwan Siskandar, Arga Aridarma, Maria Sri Kristiana Rahayu, Titah Sihdjati Riadhie, and Husin Alatas. "Non-invasive blood haemoglobin level measurement." In *AIP Conference Proceedings*, 2320, no. 1. AIP Publishing, 2021. https://doi.org/10.1063/5.0037477

[13] Raikhel, Marina. "Accuracy of noninvasive and invasive point-of-care total blood hemoglobin measurement in an outpatient setting." *Postgraduate medicine* 124, no. 4 (2012): 250-255. https://doi.org/10.3810/pgm.2012.07.2584

[14] Avcioglu, Gamze, Cemil Nural, Fatma Meriç Yilmaz, Pervin Baran, Özcan Erel, and Gülsen Yilmaz. "Comparison of noninvasive and invasive point-of-care testing methods with reference method for hemoglobin measurement." *Journal of clinical laboratory analysis* 32, no. 3 (2018): e22309. https://doi.org/10.1002/jcla.22309

[15] Shabaninejad, Hosein, Nashmil Ghadimi, Kourosh Sayehmiri, Hossein Hosseinifard, Rasoul Azarfarin, and Hasan Abolghasem Gorji. "Comparison of invasive and noninvasive blood hemoglobin measurement in the operating room: a systematic review and meta-analysis." *Journal of anesthesia* 33 (2019): 441-453. https://doi.org/10.1007/s00540-019-02629-1

[16] Acharya, Soumyadipta, Dhivya Swaminathan, Sreetama Das, Krity Kansara, Sushovan Chakraborty, Dinesh Kumar, Tony Francis, and Kiran R. Aatre. "Non-invasive estimation of hemoglobin using a multi-model stacking regressor." *IEEE journal of biomedical and health informatics* 24, no. 6 (2019): 1717-1726. https://doi.org/10.1109/JBHI.2019.2954553

[17] Yılmaz, Hakan, Burcu S. Kızılateş, Fatema Shaaban, and Ziya R. Karataş. "A novel combined deep learning methodology to non-invasively estimate hemoglobin levels in blood with high accuracy." *Medical Engineering & Physics* 108 (2022): 103891. https://doi.org/10.1016/j.medengphy.2022.103891

[18] Kavsaoğlu, A. Reşit, Kemal Polat, and Muthusamy Hariharan. "Non-invasive prediction of hemoglobin level using machine learning techniques with the PPG signal's characteristics features." *Applied Soft Computing* 37 (2015): 983-991. https://doi.org/10.1016/j.asoc.2015.04.008

[19] Sarker, Iqbal H. "Machine learning: Algorithms, real-world applications and research directions." *SN computer science* 2, no. 3 (2021): 160. https://doi.org/10.1007/s42979-021-00592-x

[20] Sumathy, B., S. Kavimullai, S. Shushmithaa, and S. Sai Anusha. "Wearable non-invasive health monitoring device for elderly using IoT." In *IOP Conference Series: Materials Science and Engineering* 1012, no. 1, p. 012011. IOP Publishing, 2021. https://doi.org/10.1088/1757-899X/1012/1/012011

[21] Dong, Ngan Thi, and Megha Khosla. "Revisiting feature selection with data complexity." In *2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE)*, pp. 211-216. IEEE, 2020. https://doi.org/10.1109/BIBE50027.2020.00042

[22] Guyon, Isabelle, and André Elisseeff. "An introduction to variable and feature selection." *Journal of machine learning research* 3, no. Mar (2003): 1157-1182.

[23] Kashef, Shima, Hossein Nezamabadi-pour, and Bahareh Nikpour. "Multilabel feature selection: A comprehensive review and guiding experiments." *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 8, no. 2 (2018): e1240. https://doi.org/10.1002/widm.1240

[24] Saarela, Mirka, and Susanne Jauhiainen. "Comparison of feature importance measures as explanations for classification models." *SN Applied Sciences* 3, no. 2 (2021): 272. https://doi.org/10.1007/s42452-021-04148-9

[25] Grant, Stuart W., Graeme L. Hickey, and Stuart J. Head. "Statistical primer: Multivariable regression considerations and pitfalls." *European Journal of Cardio-Thoracic Surgery* 55, no. 2 (2019): 179-185. https://doi.org/10.1093/ejcts/ezy403

[26] Breiman, Leo. "Random forests." *Machine learning* 45 (2001): 5-32. https://doi.org/10.1023/A:1010933404324

[27] Breiman, Leo. "Bagging predictors." *Machine learning* 24 (1996): 123-140. https://doi.org/10.1007/BF00058655

[28] Liaw, Andy, and Matthew Wiener. "Classification and regression by randomForest." *R news* 2, no. 3 (2002): 18-22.