Heart Rate Monitoring System using Max30102 Sensor and Gaussian Naive Bayes Algorithm

Beni Mustiko Aji
Faculty of Science and Technology, Yogyakarta Technology University
Siliwangi Jombor Lor, Sendangadi, Mlati, Sleman, Yogyakarta

Wahyu Sri Utami
Faculty of Science and Technology, Yogyakarta Technology University
Siliwangi Jombor Lor, Sendangadi, Mlati, Sleman, Yogyakarta

ABSTRACT
Heart health is a crucial aspect for the elderly, as it is the leading cause of death in Indonesia. However, the elderly often do not receive sufficient monitoring on a regular basis at Ngudi Makmur Integrated Health Post. Therefore, a solution is proposed to overcome this problem by introducing a heart rate monitoring system using Max30102 sensor and Gaussian Naive Bayes algorithm. The system is designed to monitor the heart rate and display the heart health condition based on the average of the measurements taken. The calibration results of the device used showed an accuracy rate of 95.8%, while testing the Gaussian Naive Bayes algorithm with the k-fold cross-validation method resulted in an accuracy of 91%. The envisioned system aims to enhance the motivation of elderly individuals at Posyandu Ngudi Makmur, encouraging them to consistently undergo routine heart health checks.

General Terms
Classification, Gaussian Naive Bayes, Heart Rate, Max30102 Sensor Monitoring, Pre-processing.

Keywords
Gaussian Naive Bayes, Heart Rate, Max30102 Sensor Monitoring.

1. INTRODUCTION
The highest number of deaths in Indonesia caused by heart disease makes heart health a crucial issue, especially for the elderly. Basic Health Research data shows a surge in the frequency of cardiovascular disease from 0.5% in 2013 to 1.5% in 2018. Cardiovascular disease can affect all ages, with the highest rates in individuals aged 65 to 75 years and above [1]. One of the relevant parameters in cardiovascular health monitoring is the average heart rate per minute. A normal average heart rate per minute ranges from 60 to 100 beats per minute in a resting state and without physical activity [2].

Regular heart health monitoring is essential to keep an eye on the heart health of the elderly. Integrated Health Post, a healthcare facility for the elderly in Indonesia, can provide a convenient option for the elderly to maintain their heart health with regular check-ups. However, Ngudi Makmur Integrated Health Post does not currently undergo heart check-ups. Every month, Ngudi Makmur Integrated Health Post only conducts blood pressure, weight, and height checks on the elderly in Krapyak, Margoagung, Seyegan, Sleman, Yogyakarta. Therefore, a tool is needed that can be used by Ngudi Makmur Integrated Health Post in monitoring the heart rate of the elderly.

Based on this description, a heart rate monitoring system using max30102 sensor and Gaussian Naive Bayes algorithm is proposed. The Naive Bayes algorithm is a statistical classifier that can be used to predict the probability of class membership while the Gaussian Naive Bayes algorithm is a variation of the Naive Bayes algorithm used to classify and analyze continuous data [3]. The Max30102 sensor is a device for measuring heart rate and calculating the average heart rate per minute with a non-invasive method that can be placed in various body locations, such as the finger. The collected data will be uploaded to Firebase via the NodeMCU ESP8266 microcontroller which enables wireless internet connectivity [4]. This data is then analyzed using the Gaussian Naive Bayes algorithm to classify heartbeat conditions. After that, the results will be displayed to monitors at the Ngudi Makmur Integrated Health Post. This research will hopefully provide an effective solution for elderly heart health monitoring, enabling them to get better care at Ngudi Makmur Integrated Health Post.

2. RELATED WORKS
The study entitled "simulation of a monitoring system for the condition of heart attack patients using Arduino and heart rate sensor" compiled by Whiliams and Sean Coonery, reviewed and analyzed information collected using Arduino and heart rate sensor as the main parameters. In the context of this study, heart rate data recorded by Arduino and sensor will be sent to smart phone devices that can detect symptoms of abnormalities or signs of arrhythmia. When such symptoms are detected, the smart phone device will issue an alert to the patient. If the alert gets no response, the smartphone device sends an emergency message. Based on the data collected, the overall results of this study concluded that the algorithm used can monitor and detect heart rate abnormalities, and this can serve as an effort to detect early symptoms of heart attack in patients during monitoring [5].

In a study entitled "Design of an Arduino Microcontroller-Based Heart Detection Tool with a Pulse Sensor" conducted by Ahmad Rizal Rinaldi, it is explained how this Android-based heart detector measures heart rate and calculates the results based on the user's age. In this study, the data was presented in the form of graphs that reflected the actual heart rate measurements that the device performed. The device allows users to record as well as transmit data regarding their heart rate. The accuracy of heart rate readings depends largely on the location of the finger used in the measurement. Testing this tool uses the mean absolute deviation (MAD) formula as a measure of accuracy that considers the magnitude of measurement errors that commonly occur. The average root-mean-square (RMS) error measurement result for applications and instruments used in heart rate and body temperature monitoring in a room without air conditioning was 4.6 for right-handed use and 9.6 for left-handed use. While the result of the right hand is 0.827 and the left hand is 0.81 [6].

The study titled "Heart rate monitor system to detect heart health levels based on the internet of things using android" was published in 2021 and written by Jarot Dian, Fujiama Diapoldo Silalahi, and Nuris Dwi Setiawan. This research explains the construction of an Internet of Things (IoT)-based heart rate monitoring system that uses Android devices and is built on the Wemos microcontroller platform. The main purpose of the system is to record and monitor the heart rate of each patient. This study was conducted at the Lebdosari Health Center, where patients often have difficulty in monitoring their heart rate. The results showed that the system was able to detect fluctuations in the patient's heart rate. The system also can provide findings and recommendations related to the user's heart rate health condition. This technology can monitor the patient's heart rate and calculate the average and maximum heart rate values [7].

3. RESEARCH METHODOLOGY

The stages of research on the development of heart rate monitoring system using the max30102 sensor and the Gaussian Naive Bayes algorithm are presented in Fig 1 below.

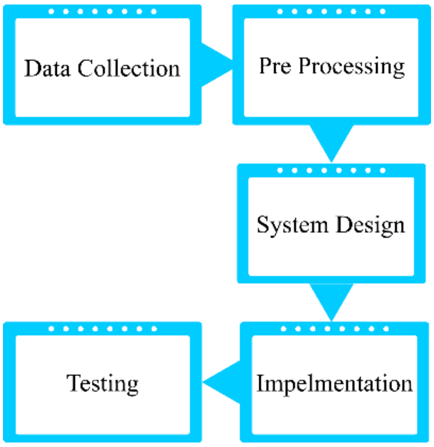

Fig 1: Research Methodology

3.1 Data Collection

Data were collected by conducting in-depth field observations through visits to Posyandu Ngudi Makmur located in Krapyak, Margoagung, Seyegan, Sleman, Yogyakarta. In addition, interviews were conducted with Posyandu administrators and heart rate measurements to obtain heart rate data.

3.2 Naive Bayes

Bayesian classifier is a statistical approach used to estimate the probability of data belonging to a particular class. One of the popular algorithms in this method is the Naive Bayes algorithm. This algorithm is a statistical and probability approach developed by an English scientist named Thomas Bayes. It uses Bayes Theorem to predict the likelihood of data belonging to a particular class based on past experience [8]. The Naive Bayes formula can be found in Equation 1.

$$P(X) = \frac{P(X|H)\,P(H)}{P(X)} \quad (1)$$

Where:

X : Unknown class data
H : Hypothesized data is a specific class
P(H|X) : Probability of hypothesis H based on condition X (posterior probability)
P(H) : Probability of hypothesis H (prior probability)
P(X|H) : Probability of X based on the condition in hypothesis H
P(X) : Probability of X

3.3 Gaussian Naive Bayes

In the case of sequential data, it is often assumed that the values associated with each category or class follow a Gaussian distribution pattern otherwise known as a normal distribution. During the data processing process, the training data is grouped into relevant classes, and then the average value (mean) and standard deviation are calculated for each of these classes [9]. Therefore, to estimate the probability of a set of consecutive data, the formula found in Equation 2 can be used.

$$P(X_i) = x_i|Y = y_j = \frac{1}{\sqrt{2\pi\sigma ij}} e^{-\frac{(x_i-\mu_{ij})^2}{2\sigma^2 ij}} \quad (2)$$

Where:

P : Chance
Xi : Value of ith attribute
Y : Class searched
yj : Subclass Y that is searched
μ : Mean, states the average of all attributes
σ : Standard deviation, expressing the variance of all attributes

3.4 Pre-Processing

Data preprocessing is a stage that serves as the first step to transform raw data into more structured, clean, and consistent information [10]. The main purpose of data preprocessing in the context of this study is to prepare the data needed for further analysis, in particular the use of Max30102 sensor and Gaussian Naive Bayes algorithms in heart rate monitoring in the elderly at the Ngudi Makmur Integrated Health Post so that the raw data can be used for further processing steps.

3.5 System Design

System design in the heart rate monitoring system using the max30102 sensor is done by analyzing the system first to find out what process will be made next. The next stage is to create a system design that describes the flow of the system such as flowchart design.

3.6 Implementation

After the development of the system design, the heart rate monitoring system using max30102 sensor and Gaussian Naive Bayes algorithm was implemented by writing code using java programming language for the system and connecting it to the database.

3.7 Testing

Evaluation of the results of heart rate monitoring using the max30102 sensor was carried out by testing the heart rate calibration and testing the accuracy of the model using the k-fold cross validation method produced in the calculation of the Gaussian Naive Bayes algorithm.

4. RESULTS AND DISCUSSION

The test results of the Android-based heart rate monitoring system using the max30102 sensor and the Gaussian Naive Bayes algorithm were carried out by taking elderly heart rate data at the Ngudi Makmur Integrated Health Post and processed using the Gaussian Naive Bayes algorithm. The results of the accuracy level of calculating the average heart rate using the max30102 sensor using heart rate calibration measurement techniques and accuracy measurements with the k-fold cross validation method on the calculation of the Gaussian Naive Bayes algorithm.

4.1 Data Collection

Heart rate calculation by experts is done using a stethoscope attached to the chest and heart rate measurement using the max30102 sensor by placing the tip of the thumb on the max30102 sensor as shown in Fig 2 below.

Fig 2: Data Collection Process

The following are the results of heart rate calculations in the elderly conducted by experts and max30102 sensor shown in table 1.

Table 1. Heart Rate Data

Number	Age (Year)	Gender	max30102 Sensor	Expert
1	61	P	78	74
2	60	L	44	50
3	64	P	49	53
4	62	P	98	96
5	63	L	70	66
….	….	….	….	….
294	65	P	106	110
295	66	L	108	102
296	68	P	118	116
297	61	P	65	71
298	67	L	57	61

Data in this study amounted to 298 with data collection carried out regularly. The data will be calculated for accuracy by heart rate calibration to determine the level of accuracy of the tool used.

4.2 Pre-processing

The data used in this study was subjected to an encoding process at the data pre-processing stage. Encoding is an activity carried out by the source to translate his thoughts and ideas into a form that can be accepted by the recipient's senses [11] by changing existing features. For example, in the gender feature, data (P) which means Female is changed to the number one (1), while data (L) which means male is changed to the number zero (0). The following calculation results can be seen in table 2.

Table 2. Heart Rate Data Encoding Results

Number	Age (Years)	Gender	Measurement Results
1	61	1	78
2	60	0	44
3	64	1	49
4	62	1	98
5	63	0	78
….	….	….	….
294	65	1	106
295	66	0	108
296	68	1	118
297	61	1	65
298	67	0	57

4.3 System Design

In this study, system design includes creating flowchart diagrams that visually illustrate the workflow of the system. In Bahasa Indonesia, flowchart referred to as flowchart is a graphical representation of the logic and organization of a program [12]. This diagram is a graphical representation of the logical processes involved in a computer program or operating system [13]. The system flowchart is shown in Fig 3.

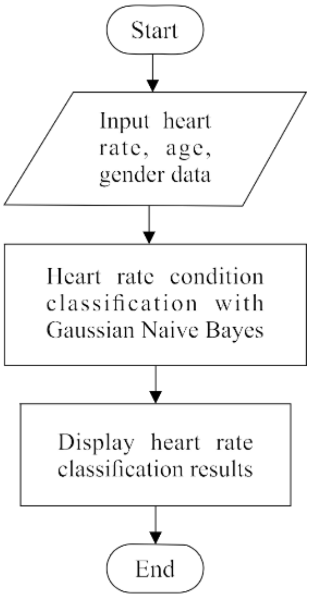

Fig 3: Flowchart

4.4 Implementation

The implementation stage is the process of converting or implementing system design into tangible form. This process ensures that the system implemented conforms to requirements and expectations by designing, creating, testing, and integrating various parts of the system. The implementation of a heart rate monitoring system using the max30102 sensor and the Gaussian Naive Bayes algorithm will involve several important stages.

4.5 Testing

4.5.1 Heart Rate Calibration Accuracy Testing

At the testing stage, in addition to system testing, there is testing the level of calculation accuracy of the system using the heart rate calibration test method. Calibration is common practice to compare one measurement standard with another. These measurements are generally recognized for national and international accreditation. This means that relevant indicators will be met in the context of the prevailing conditions or system. In this study, calibration is interpreted as a comparison between heart rate measured with estimated heart rate using max30102 sensor [14].

Data collected through stethoscope readings were used to calibrate the heart rate monitor and its findings were compared to max30102 to determine the difference in the calculation of the average heart rate. The heart rate calibration formula looks at equation 3.

$$Accuracy\ value(\%) = 100\% - \left|\frac{Difference}{Commercial\ tools}\right| \times 100 \quad (3)$$

The Max30102's heart rate calibration was calculated manually using data from many people in a resting (sitting) state. Table 3 displays the nonconformity and accuracy values calculated using the collected data.

Table 3. Results of Average Heart Rate Accuracy

Number	Age (Year)	Gender	Max30102	Expert	Difference	Accuracy
1	61	P	78	74	4	96%
2	60	L	44	50	6	94%
3	64	P	49	53	4	96%
4	62	P	98	96	2	98%
5	63	L	70	66	4	96%
….	….	….	….	….	…	….
294	65	P	106	110	4	96%
295	66	L	108	102	6	94%
296	68	P	118	116	2	98%
297	61	P	65	71	6	94%
298	67	L	57	61	4	96%
Average accuracy						95,8%

After collecting data using the max30102 sensor, the data from the max30102 sensor was compared with the expert's measurement. In Table 3, the difference between the max30102 sensor and the expert's measurement shows a sensor accuracy of 95.8% in estimating the average heart rate.

4.5.2 Gaussian Naive Bayes Algorithm Testing

In evaluating the Gaussian Naive Bayes algorithm, Fig 4. illustrates the data distribution in the data set used in this study.

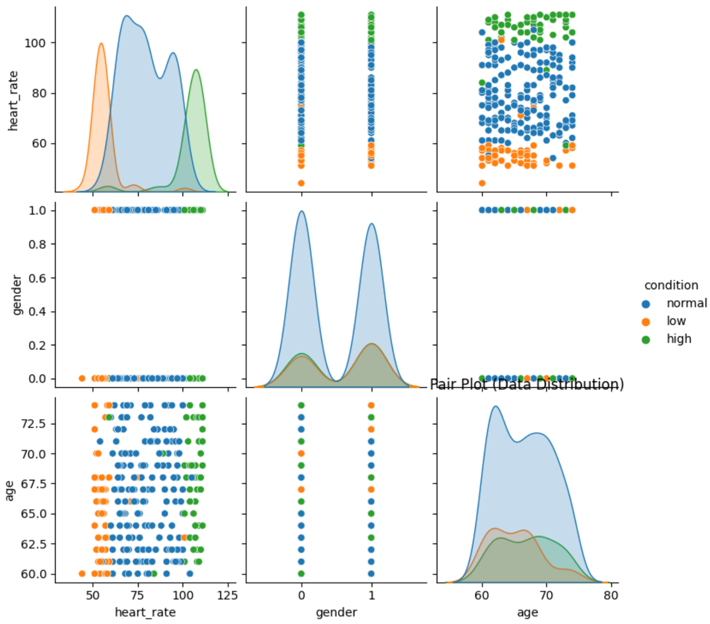

Fig 4: Data Distribution

At this stage, the calculation of the mean and standard deviation of each feature (heart rate, gender and age) in all existing classes (high, normal and low) is carried out. Here are the results of mean and standard deviation with the naïve bayes algorithm in table 4.

Table 4. Results of Mean and Standard Deviation Calculation

Class	Feature	Mean	Standard Deviation
High	Heart rate	103.64	11.34
High	Gender	0.52	0.51
High	Age	66.73	4.4
Normal	Heart rate	80.02	12.74
Normal	Gender	0.49	0.5
Normal	Age	66.54	4.38
Low	Heart rate	55.7	8.12
Low	Gender	0.55	0.5
Low	Age	64.55	3.78

At this stage the researcher calculated the probability of Prior arising from the number of data sets totaling 298 data. Here is table 5 of prior probabilities on the application of the Gaussian Naive Bayes algorithm.

Table 5. Results of the calculation of the probability of prior classes appearing

Number	Class	Prior class probability
1	Normal	0.63
2	High	0.19
3	Low	0.19

After finding the prior class probability in this study, here is an example of a detailed calculation that can predict heart rate conditions using the Gaussian Naive Bayes formula using data:

Given data point:

Heart Rate (X) = 106 bpm

Gender (G) = Female (1)

Age (A) = 65 years old

calculating the posterior probability for each class and predicting the label with the maximum probability.

Prior:

P(High) = 0.19

P(Normal) = 0.63

P(Low) = 0.19

Average value and Standard Deviation for each feature in each class:

High:

$\mu X = 103.64, \sigma X = 11.34$

$\mu G = 0.52, \sigma G = 0.51$

$\mu A = 66.73, \sigma A = 4.4$

Normal:

$\mu X = 80.02, \sigma X = 12.74$

$\mu G = 0.49, \sigma G = 0.5$

$\mu A = 66.54, \sigma A = 4.38$

Low:

$\mu X = 55.7, \sigma X = 8.12$

$\mu G = 0.55, \sigma G = 0.5$

$\mu A = 64.55, \sigma A = 3.78$

for the calculation of each posterior can be seen as:

P(Height | 106, 1, 65)

$= (1/\sqrt{(2\pi*11.34)}) * \exp(-(106-103.64)2 / (2*11.342))$

$* (1/\sqrt{(2\pi*0.51)}) * \exp(-(1-0.52)2 / (2*0.512))$

$* (1/\sqrt{(2\pi*4.4)}) * \exp(-(65-66.73)2 / (2*4.42))$

= 0.16

This is the result of the class calculation obtained

P(Normal | 106, 1, 65) = 0.034

P(Low | 106, 1, 65) = 7.9e-06

The highest probability is for the High class. Hence, this calculation data point is classified as "High" heart rate condition using Naive Bayes.

At this stage of calculation, researchers use the k-fold cross-validation method. In cross validation, there are two popular approaches to evaluating algorithm performance, namely k-fold cross validation and leave-one-out cross validation. When the amount of data is large, k-fold cross validation should be used to estimate the accuracy of the data [15]. Fig 5 shows the accuracy measurement result with k value equal to 5.

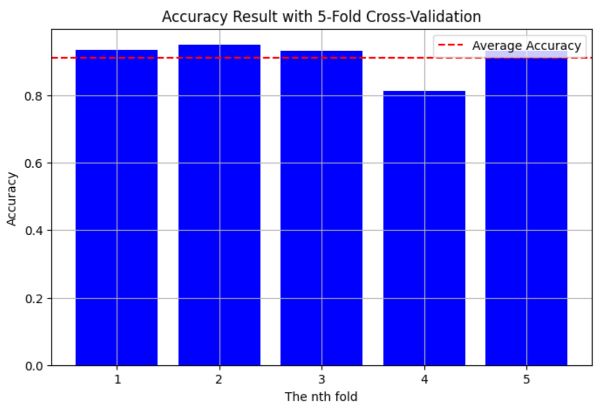

Fig 5: Accuracy results with K-Fold Cross-Validation

Based on the calculations carried out, the accuracy rate of using the Gaussian Naive Bayes algorithm in the heart rate monitoring system using the max30102 sensor at Ngudi Makmur Integrated Health Post is 91%.

5. CONCLUSION

Based on the results of this study, it can be concluded that the heart rate monitoring system using max30102 sensor and Gaussian Naive Bayes algorithm in terms of functionality with black box testing shows that the implementation of the system design runs well. Heart rate calibration testing shows an accuracy of 95.8%. The accuracy of the Gaussian Naive Bayes algorithm measured using the k-fold cross validation method resulted in an accuracy of 91%. The heart rate monitoring system using the max30102 sensor and the gaussian naïve bayes algorithm can be applied to the elderly at Posyandu Ngudi Makmur to help check well and can be applied sustainably.

This research still has shortcomings, one of which is the algorithm used. The gaussian naive bayes algorithm has limitations in handling categorical features and dependence on the amount of training data so further research is recommended to use a fuzzy algorithm that can provide maximum measurement results because it has the advantage of processing fuzzy data well, such as noisy or erratic sensor data. This can help in producing more stable and accurate measurement accuracy so that the determination of heart rate conditions becomes better.

6. ACKNOWLEDGMENTS

This research still has shortcomings, one of which is the algorithm used. The gaussian naive bayes algorithm has limitations in handling categorical features and dependence on the amount of training data so further research is recommended to use a fuzzy algorithm that can provide maximum measurement results because it has the advantage of processing fuzzy data well, such as noisy or erratic sensor data. This can help in producing more stable and accurate measurement accuracy so that the determination of heart rate conditions becomes better.

7. REFERENCES

[1] Kementrian Kesehatan Republik Indonesia, “Hasil Utama Riskesdas 2018,” Jakarta, 2018.

[2] L. Herawati et al., Personal Trainer Guide, 1st ed. Departemen Ilmu Faal Fakultas Kedokteran Universitas Airlangga, 2019.

[3] K. Hajer, A. Dhahir, and J. A. Al-tuwairijari, “Cancer Classification Using Gaussian Naive Bayes Algorithm,” in International Engineering Conference (IEC), Erbil, Iraq: IEEE, Jun. 2019, pp. 165–166.

[4] D. N. Ilham, Hardisal, and R. A. Candra, Monitoring dan Stimulasi Detak Jantung dengan Murotal Al-Qur’an Berbasis Internet of Things (IoT). CV Jejak, anggota IKAPI, 2020.

[5] Whiliams and S. C. Sumarta, “Simulasi Sistem Monitoring Kondisi Pasien Serangan Jantung Menggunakan Arduino Dan Sensor Detak Jantung,” Jurnal Tematika, vol. VOL. 10, NO. 1, pp. 46–49, Mar. 2022.

[6] A. R. Rinaldi, “Rancang Bangun Alat Deteksi Jantung Berbasis Mikrokontroler Arduino Dengan Pulse Sensor,” Forum Pendidikan Tinggi Teknik Elektro Indonesia Regional VII, pp. 374–376, 2021.

[7] J. Dian, F. D. Silalahi, and N. D. Setiawan, "Sistem Monitoring Detak Jantung Untuk Mendeteksi Tingkat Kesehatan Jantung Berbasis Internet Of Things Menggunakan Android," Jurnal JUPITER, vol. Vol. 13 No. 2, pp. 69–72, Oct. 2021.

[8] R. Mayasari, N. Nafisah, and A. Natuzzuharryyah, "Klasifikasi Tingkat Kepuasan Mahasiswa Terhadap Pembelajaran Secara Daring Menggunakan Algoritma Naive Bayes," JISka (Jurnal Informatika Sunan Kalijaga), vol. Vol. 6, No. 3, p. 163, Jun. 2021.

[9] H. Kamel, D. Abdulah, and J. M. Al-Tuwaijari, "Cancer Classification Using Gaussian Naive Bayes Algorithm," in International Engineering Conference (IEC), Erbil, Iraq: IEEE, Jun. 2019, pp. 166–166.

[10] Y. Findawati and M. A. Rosid, Buku Ajar Text Mining, Pertama. Sidoarjo: UMSIDA PRESS, 2020.

[11] A. Géron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, Second. O'Reilly Media, Inc., 2019.

[12] U. Indahyant and Y. Rahmawati, Algoritma dan Pemrograman Dalam bahasa c++, Pertama. Sidoarjo: UMSIDA Press, 2020.

[13] D. Indra Ava, Logika dan Algoritma Untuk Merancang Aplikasi Komputer , 1st ed. Semarang: Yayasan Prima Agus Teknik, 2021.

[14] F. Rouessac and A. Rouessac, Analysist Modern Intrumentation Method Techques. West Sussex: Willy & Sons, 2007.

[15] R. Kumar, Machine Learning Quick Reference. Packt Publishing, 2019.