http://jhmi.sums.ac.ir
Health Man & Info Sci 2023, 10(2), 87–92

HMIS
Health Management and Information Science

Original Article

Classification of Maternal Emergencies Using Gaussian Naive Bayes to Speed up the Patient's Triage Process

Evvin Faristasari[1*], Igi Ardiyanto[2], Eugenius Phyoway Ganap[3]

[1]Master of Biomedical Engineering Graduate School of Gadjah Mada University, Yogyakarta, 555281, Indonesia
[2]Department of Electrical Engineering and Information Technology, Faculty of Engineering of Gadjah Mada University, Yogyakarta 55281 Indonesia
[3]Department of Obstetrics and Gynecology Faculty of Medicine, Public Health, and Nursing of Gadjah Mada University, Yogyakarta 55281 Indonesian

Abstract
Introduction: Labor is the most important process in every woman's pregnancy. This requires optimal handling of various parties until labor takes place smoothly. The purpose of the study is to determine the triage classification of labor referral patients in hospitals using Gaussian Naive Bayes as the final model.
Methods: This study used 90 data, each consisting of 15 parameters which are divided into two categorical data types: 9 data and 6 continuous data types. Two treatments were used in this study, namely Gaussian Naive Bayes (first) using the independence assumption on all parameters, and Categorical Naive Bayes for categorical data types, and Gaussian Naive Bayes for continuous data types. These two types of data were combined using Gaussian Naive Bayes as the final model. The data went through a preprocessing stage, stratified cross-validation; then, we used the method of Naive Bayes according to the data type and continued for the final stage classification using Gaussian Naive Bayes.
Results: The results of the first treatment had an accuracy of 91%, recall of 97%, precision of 64%, and F1-score of 73%. Also, the second treatment had an accuracy of 96%, recall of 88%, precision of 86% and F1-score of 86%. The treatment of different data types had a difference in the final results compared to the treatment of the same data type.
Conclusion: The diversity of data types is best handled according to the model used.
Keywords: Labor referral, Triage classification, Gaussian naive bayes

Article History:
Received: 18 March 2023
Accepted: 06 June 2023

Please cite this paper as:
Faristasari E, Ardiyanto I, Ganap EP. Classification of Maternal Emergencies Using Gaussian Naive Bayes to Speed up the Patient's Triage Process. Health Man & Info Sci. 2023; 10(2): 87-92. doi: 10.30476/JHMI.2024.100872.1192.

***Correspondence to:**
Evvin Faristasari,
Master of Biomedical Engineering Graduate School of Gadjah Mada University, Yogyakarta, 555281, Indonesia
Email: evvin.faristasari@mail.ugm.ac.id

Introduction

Every pregnant woman has potential risks during pregnancy, labor, and puerperium (1). Efficient care for pregnant women with obstetric emergencies is needed to improve maternal and fetal health (2). In the Emergency Department (ED), a rapid assessment is needed to overcome the patient's emergency condition by triaging quickly and precisely, but the limited information obtained during triage can cause inappropriate decision making (3). Triage is a process of classifying patients based on condition assessment which is divided into 5 categories, namely red (immediate, life-threatening), orange/orange (emergency, condition can decline rapidly), yellow (urgent treatment is needed before other complications arise), green (less urgent so we can wait until other emergencies are under control), and blue (not urgent so treatment can be postponed until conditions calm down) (4). Triage is carried out continuously from the referral point of origin (before arriving at the hospital) by preparing transportation and sequencing of patients until the patient arrives at the hospital (5). Triage aims to optimize the performance of medical personnel in the emergency room and is an effort to reduce mortality by moving patients according to general status and examination results (5). The primary role of obstetric triage is to evaluate labor and for emergency management of the patient by assessing the maternal patient's Vital Signs (VTS), Fetal Heart Rate (FHR), pain score, signs of labor, and risk score (6). Several factors affecting referral quality were discussed by previous researchers (7), explaining the importance of communication during referrals and (8) indicating that insurance affects prenatal care. During the referral process, there are several processes that can affect the patient's condition. Phyo Referral Score (PRS) is a tool that has a validity test of 93.3%, sensitivity of 72.7%, and specificity of 88.9%, so it can be used to assess the quality of referral of maternal patients by taking into account several parameters such as age (mother's age), cause (cause of referral), staff (assistance by medical personnel), procedure (referral procedure), phone (referrer information), ambulance (ambulance when referring), insurance (insurance used by the patient), area, and district (distance to the referral place) (9).

Data mining is a method used to learn patterns from data sets using several methods such as statistics, machine learning, mathematics, or artificial intelligence (10). The role of artificial intelligence has been evaluated using machine learning (11). A collection of health data may be able to generate useful information for public health improvement by applying machine learning (12). There are three categories in machine learning, they are supervised learning, unsupervised learning, and reinforcement (13). Naive Bayes Classifier (NBC) is a simple machine learning that can perform tasks such as predicting medical diagnoses, spam filtering, and forecasting (14). A study (15) entitled "Towards applying the Internet of things and machine learning for the risk prediction of COVID-19 in pandemic situations using Naive Bayes classifier for improving accuracy" explains that NBC has good performance for prediction in machine learning. The problem raised in this study is that due to limited data conditions, maternal triage prediction is carried out using the Naive Bayes method to determine the performance results of the model used. This study aimed to determine how to apply Gaussian Naive Bayes as the final model to predict the triage of maternal patients with 15 parameters from the PRS score and the results of the patient's examination when arriving at the Emergency Department of Comprehensive Emergency Obstetric Neonatal Services and to determine the classification results.

Materials and Methods

This study was conducted at Wates Regional General Hospital, Kulon Progo Regency, Yogyakarta Special Region. Wates Regional General Hospital is a Regency referral hospital that has Comprehensive Emergency Obstetric Neonatal Services. The research was conducted in November 2021. The population in this study were maternal referral patients from January to March 2020. The sample used consisted of maternal referral patient data taken from January to March 2020.

The sampling method used in this research was purposive sampling, samples are selected based on predetermined criteria. The sample size was 90 ssubjects for the training data.

a. Inclusion criteria:

1. Maternal referral patients with gestational age over 36 weeks (350/7-6/7) (16)
2. Maternal referral patients without indications of COVID-19 in the emergency room from January to March 2020

b. Exclusion criteria:

1. Maternal referral patients with gestational age under 36 weeks
2. Post-natal care referral patients
3. Maternal referral patients with incomplete medical record data
4. Maternal referral patients with COVID-19 indications

Research Variables

The variables in this study consisted of the independent variable for features (x) and the dependent variable for class (y). The independent variable consisted of 15 parameters/attributes taken from the PRS score (9) and the results of the maternal patient examination (17, 18) which consisted of age, cause, staff, stabilization (stabilization received by the patient during the referral process), phone (referrer's phone), ambulance (referrer's ambulance), insurance (insurance owned by the patient), area (area), distance (referral distance to the hospital / referral place), systole (systole in the patient's blood pressure), diastole (diastole in the patient's blood pressure), Respiratory Rate (patient RR), temperature (patient temperature), pulse (patient pulse), and FHR (Fetal Heart Rate) with the following explanation in Table 1. The dependent variable shown in Table 2 consisted of 3 classes, namely orange class, yellow class, and blue class.

The independent variable had 9 discrete data types and 6 continuous data types. Discrete data started from age to distance, while continuous data started from systole, dyastole, FHR. The raw dataset was processed first by imputation in this study using the median for imputation; then, after imputation, assuming discrete independent data and cross validation was performed with 5-level cross validation due to unbalanced data conditions.

Research Stages

Data were collected through interviews with hospital staff, observation of maternal documents, and literature review. Next, preprocessing was carried out by performing data preparation (cleaning) and imputation and then using stratified cross-validation with 5 folds with imbalance dataset.

Table 1: Dependent variables

No.	Triage	Code
1.	Blue	3
2.	Yellow	2
3.	Orange	1

Table 2: Independent variables

Feature (x_i)	Code of Feature Values			Type Data
	0	1	2	
Age	<20 or >35	20–35		Discrete
Cause	Avoidable	Unavoidable		Discrete
Staff	Non Medic	Medic		Discrete
Stabilization	Without stabilization	Stabilization		Discrete
Phoned	Without Phoned	Phoned		Discrete
Ambulance	Non Ambulance	Ambulance		Discrete
Insurance	Non Insurance	Insurance		Discrete
Area	Different district	Same district		Discrete
Distance	>30 km	10 – 30 km	10>	Discrete
Systole	(mmHg)			Continuous
Diastole	(mmHg)			Continuous
Respiratory Rate (RR)	(x/minute)			Continuous
Temperature	(°C)			Continuous
Pulse	(x/minute)			Continuous

Application of Naive Bayes Model

Naive Bayes is one type of classification in data mining that uses Bayesian rules and probability theorems (19). In this research, the data were of discrete and continuous types so that it uses several methods from Naïve Bayes with calculations using data independence (continuous or discrete) and combining both types of data. The Bayes theorem equation can be seen in equation 1 as follows:

$$P(X) = \frac{P(Y)P(Y)}{P(X)} \quad (1)$$

Where, X is the variable (x1, x2, x3, ... , xn), Y is the class (orange, yellow, or blue), P(X|Y) is the probability, P(Y) is the prior probability, and P(X) is the marginal probability (20). In this research, we used several methods of Naive Bayes including:

1. Gaussian Naive Bayes

In Gaussian Naive Bayes continuous attributes consisted of numerical data (21). All type data in each parameter used Gaussian Naive Bayes with equation :

$$P(x_i \mid y = y') = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{\left(-\frac{(x_i - \mu_x)^2}{2\sigma_x^2}\right)} \quad (2)$$

with σ as standard deviation and μ as the mean.

2. Categorical Naive Bayes (NB) *and Gaussian Naive Bayes with Gaussian Naive Bayes as a final model*

a. Categorical NB

CategoricalNB was used for categorically distributed data; the probability of category t in feature i given class c had the following equation:

$$P(x_i = t \mid y = c; \alpha) = \frac{N_{tic} + \alpha}{N_c + \alpha n_i} \quad (3)$$

It is assumed that each feature described by the index i has a categorical distribution. For each feature i in the training set X, CategoricalNB assumes a categorical distribution for each feature i of X conditioned on class y. The set of indices on the sample is defined as $J = \{1, \ldots m\}$ with m being the number of samples.

$$N_{tic} = |\{j \epsilon J \mid x_{ij} = t, y_j = c\}|$$

is the category of how many times t appears in sample x_i, which belongs to class c, $N_c = |\{j \epsilon J \mid y_j = c\}|$ is the number of samples with class c, α is smoothing parameter, and n_i is the number of available categories of feature i.

In this research, 9 feature data with type data discrete were calculated using CategoricalNB (22).

b. Gaussian Naive Bayes

For 6 features that had a continuous data type, Gaussian Naive Bayes (GaussianNB) calculation was used.

c. Final Model

The results of the CategoricalNB and GaussianNB calculations, according to the data type, obtained the results in the form of predicted probabilities. This predicted probability will be later calculated using the final model ofGaussian Naive Bayes.

Results

Preprocessing

In the preprocessing stage, there were several features that had missing values, namely in the staff, stabilization, ambulance, insurance, and stabilization columns, as shown in Table 3.

Table 3: Missing Values

Feature	Missing Values
Age	0
Cause	0
Staff	1
Stabilization	11
Phone	0
Ambulance	7
Insurance	2
Area	0
Distance	0
Systole	0
Diastole	0
Respiration Rate	0
Temperature	0
Pulse	0
Fetal Heart Rate	0

Validation

Data proportion was the next stage to be analyzed. There was an imbalance in the data, as shown in Figure 1, with three classes: Code 1 for orange class, code 2 for yellow class, and code 3 for blue class. A strategy to confirm that each class proportion receives equal treatment in this study is to use 5-fold stratified cross-validation which is displayed in Figure 1.

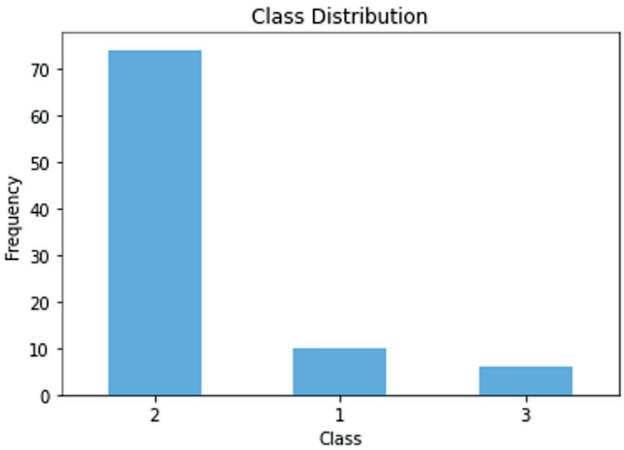

Figure 1: Class distribution

Classification Algorithm Performance Measurement

The dataset of 90 maternal referral patient data was applied to Gaussian Naive Bayes with 2 treatments; then, performance assessment was carried out by performing accuracy, recall, precision, and F1-score; the results are shown in Table 4.

Discussion

The main focus of this study was to determine the results of triage classification of labor referral patients using Gaussian Naive Bayes with different treatments based on the type of data on the available features according to the conditions of maternal emergencies. Sample selection was carried out based on patient inclusion and exclusion criteria so that the data was homogeneous from the start; this is justified by (23) which discusses that sample selection based on inclusion and exclusion criteria has an impact on the external validity of the study.

Based on the research conducted by (9), several parameters can be assessed in cases of referral of maternal emergencies to the hospital, including age, cause, staff, stabilization (stabilization received by the patient during the referral process), phone (referrer's phone), ambulance (referrer's ambulance), insurance (insurance owned by the patient), area (area), distance (referral distance to the hospital/referral place). Other parameters can be observed from the condition of the mother during labor and the vital signs that have contributed to the smooth process of labor (17, 18); these parameters include the referral patient's condition, such as systole (systole in the patient's blood pressure), diastole (diastole in the patient's blood pressure), respiration rate during labor, and should be known to anticipate pathological events during labor, temperature (patient temperature), pulse (patient pulse), and FHR (Fetal Heart Rate).

Data has various types such as discrete data,

Table 4: Classification result

Value	Methods	
	GNB (first treatment)	GNB Combine (second treatment)
Accuracy	91%	96%
Recall	97%	88%
Precision	64%	85%
F1-score	73%	86%

categorical data, continuous data, and others. The treatment of data in this study had different results in the final stage of classification. In the first treatment, different types of data on the parameters were treated independently and processed using one model, namely Gaussian Naive Bayes, which has an accuracy value of 91% and F1-score of 73%. In the second treatment with 9 data using Categorical Naive Bayes, 5 data using Gaussian Naive Bayes, and the last model using Gaussian Naive Bayes from the results of combining predicted probabilities of both types of data resulted in an accuracy value of 96% and F1- score of 86%. The treatment of this data is very interesting although both use Gaussian Naive Bayes as the final model. It is due to the varying conditions of the data that we are interested in giving this treatment.

Conclusion

Medical data has a wealth of information that can be useful for the healthcare field. This wealth of information is available in various types of data that can certainly help the performance of medical personnel in making decisions. In this study, the data focused on maternal referrals following the inclusion and exclusion criteria. Different results were obtained between the data directly processed using the final Gaussian Naive Bayes model and data processed first with several models, namely Categorical Naive Bayes and Gaussian Naive Bayes according to data type and the final model using Gaussian Naive Bayes. This shows that data type can affect the performance of machine learning.

Acknowledgement

We would like to thank all the hospital staff who assisted us during this study.

Ethics Approval

This research was approved by the Health Research Ethics Committee of the Wates Regional General Hospital with Number, No.KEPK/079/RS/X/2021.

Financial Support

No funding was received for this research.

Conflict of Interest: None declared.

References

1. Holness N. High-risk pregnancy. *Nursing Clinics*. 2018;53(2):241-51. doi: 10.1016/j.cnur.2018.01.010.
2. Sert ZS, Sert ET, Kokulu K. Predictors of obstetric complications following traumatic injuries in pregnancy. *Am J Emerg Med*. 2021;45:124-8. doi: 10.1016/j.ajem.2021.02.056.
3. Huang YB, Jiang YF, Li YR, Yuan YS, Clutter P, Chi FL, et al. A Predictive Model Development of Hospital Admission during Triage in a Chinese Ear, Nose, and Throat Emergency Department. *J Emerg Nurs*. 2021;47(6):914-24. doi: 10.1016/j.jen.2021.04.012.
4. Hoffmann E, Wilburn-Wren K, Dillon SJ, Barahona A, McIntire DD, Nelson DB. Impact of implementation of the Maternal-Fetal Triage Index on patients presenting with severe hypertension. *Am J Obstet Gynecol*. 2022;227(3):521 e1- e8. doi: 10.1016/j.ajog.2022.06.006.
5. Demirel ME, Ali IH, Bogan M. Emergency service experience following the terrorist attack in Mogadishu, 14 October 2017, a scene of lay rescuer triage. *Am J Emerg Med*. 2021;40:6-10. doi: 10.1016/j.ajem.2020.12.005.
6. Kodama S, Mokhtari NB, Iqbal SN, Kawakita T. Evaluation of the Maternal-Fetal Triage Index in a tertiary care labor and delivery unit. *Am J Obstet Gynecol MFM*. 2021;3(4):100351. doi: 10.1016/j.ajogmf.2021.100351.
7. Ameyaw EK, Amoah RM, Njue C, Tran NT, Dawson A. Women's experiences and satisfaction with maternal referral service in Northern Ghana: A qualitative inquiry. *Midwifery*. 2021;101:103065. doi: 10.1016/j.midw.2021.103065.
8. Eliason EL. Adoption of Medicaid Expansion Is Associated with Lower Maternal Mortality. *Womens Health Issues*. 2020;30(3):147-52. doi: 10.1016/j.whi.2020.01.005.
9. Ganap EP, Hakimi M, Emilia O, editors. Assessment of Maternal Referral Systems in Dr. Sardjito General Hospital Using Phyo Referral Score: A Single-Center Preliminary Study. Atlantis: 4th International Symposium on Health Research (ISHR 2019); 2020. p. 238-240. doi: 10.2991/ahsr.k.200215.045.
10. Gordan M, Sabbagh-Yazdi S-R, Ismail Z, Ghaedi K, Carroll P, McCrum D, et al. State-of-the-art review on advancements of data mining in structural health monitoring. *Measurement*. 2022;193:110939. doi: 10.1016/j.measurement.2022.110939.
11. Mintz Y, Brodie R. Introduction to artificial intelligence in medicine. *Minim Invasive Ther Allied Technol*. 2019;28(2):73-81. doi: 10.1080/13645706.2019.1575882.
12. Chauhan C, Gullapalli RR. Ethics of AI in Pathology: Current Paradigms and Emerging Issues. *Am J Pathol*. 2021;191(10):1673-83. doi: 10.1016/j.ajpath.2021.06.011.

13. Rashidi HH, Tran NK, Betts EV, Howell LP, Green R. Artificial Intelligence and Machine Learning in Pathology: The Present Landscape of Supervised Methods. *Acad Pathol*. 2019;6:2374289519873088. doi: 10.1177/2374289519873088.
14. Shaban WM, Rabie AH, Saleh AI, Abo-Elsoud MA. Accurate detection of COVID-19 patients based on distance biased Naive Bayes (DBNB) classification strategy. *Pattern Recognit*. 2021;119:108110. doi: 10.1016/j.patcog.2021.108110.
15. Deepa N, Sathya Priya J, Devi T. Towards applying internet of things and machine learning for the risk prediction of COVID-19 in pandemic situation using Naive Bayes classifier for improving accuracy. *Mater Today Proc*. 2022;62:4795-9. doi: 10.1016/j.matpr.2022.03.345.
16. Spong CY, Saade G, Mercer B, D'Alton M, Blackwell S, Kilpatrick SJ. Concept of Gestational Age in "Completed Weeks": Lost in Translation. *Obstetrics & Gynecology*. 2012;119(1):184. doi: 10.1097/AOG.0b013e31823f012c.
17. Pandey D, Bharti R, Dabral A, Khanam Z. Impact of WHO Labor Care Guide on reducing cesarean sections at a tertiary center: an open-label randomized controlled trial. *AJOG Glob Rep*. 2022;2(3):100075. doi: 10.1016/j.xagr.2022.100075.
18. Annamraju H, Robertson K, Mackillop L. Respiratory disease in pregnancy. *Obstetrics, Gynaecology & Reproductive Medicine*. 2020;30(5):152-5. doi: 10.1016/j.ogrm.2020.02.007.
19. Jahromi AH, Taheri M, editors. A non-parametric mixture of Gaussian naive Bayes classifiers based on local independent features. 2017 Artificial intelligence and signal processing conference (AISP); 2017. p. 209-12. doi: 10.1109/AISP.2017.8324083.
20. Zhang H, Liu CT, Mao J, Shen C, Xie RL, Mu B. Development of novel in silico prediction model for drug-induced ototoxicity by using naive Bayes classifier approach. *Toxicol In Vitro*. 2020;65:104812. doi: 10.1016/j.tiv.2020.104812.
21. Islam R, Devnath MK, Samad MD, Al Kadry SMJ. GGNB: Graph-based Gaussian naive Bayes intrusion detection system for CAN bus. *Vehicular Communications*. 2022;33:100442. doi: 10.1016/j.vehcom.2021.100442.
22. Scikit-learn [Internet]. Multinomial Naive Bayes. [c2020]. https://scikit-learn.org/stable/modules/naive_bayes.html#categorical-naive-bayes
23. Patino CM, Ferreira JC. Inclusion and exclusion criteria in research studies: definitions and why they matter. *J Bras Pneumol*. 2018;44(2):84. doi: 10.1590/s1806-37562018000000088.