



Asthma Risk Classification in Pediatrics Using Gaussian Naive Bayes

Dyg Khayrunsalihaty Bariyyah binti Abang Othman*, Azarina binti Azman
Department of Electrical Engineering, Politeknik Kuching Sarawak, Kuching, 93050, Malaysia.

Article Info	Abstract
Article history: Received Jun 12 th , 2025 Revised Aug 4 th , 2026 Accepted Aug 16 th , 2026 Published Sep 30 th , 2026	Pediatric asthma remains a significant public health concern, often complicated by challenges in early detection and consistent monitoring. Assessing asthma risk in pediatric patients, particularly those under five years of age, is challenging because traditional diagnostic tests often require patient cooperation and precise breathing maneuvers, which are difficult for young children to perform. Caregivers may also have varying levels of knowledge and different perceptions regarding the severity of asthma risk. To enable rapid and non-invasive triage, this study proposes an asthma risk classification approach using a machine learning model based on the Gaussian Naive Bayes (GNB) algorithm. The model was trained on synthetic data designed to approximate real-world physiological patterns, with heart rate and peripheral blood oxygen saturation (SpO₂) serving as the key input features and asthma risk levels (high, moderate, low) as the target classes. The trained GNB model achieved an overall classification accuracy of 86% on the test set. Evaluation using the F1-score revealed 0.94 for the low-risk class (Precision: 0.92, Recall: 0.96), and 0.78 for the high-risk (Precision: 0.88, Recall: 0.70), while the moderate-risk class achieved an F1-score of 0.77 (Precision: 0.75, Recall: 0.80). GNB also the Decision Tree (81.00%) and Random Forest (81.00%) baselines and slightly exceeded the performance of the Support Vector Machine (85.00%). Furthermore, the evaluation of model behavior showed that accuracy gap between the training and testing results remained below 0.03 toward the end of the learning curve, indicating limited overfitting on the synthetic dataset.
Index Terms: Machine learning Gaussian Naive Bayes Risk Classification Pediatric Asthma	

This is an open access article under the [CC BY-NC-ND 4.0 license](https://creativecommons.org/licenses/by-nc-nd/4.0/).



*Corresponding Author: khayrunsalihaty@poliku.edu.my

I. INTRODUCTION

Artificial Intelligence (AI) and its subfield, Machine Learning (ML), are revolutionizing the landscape of medical technology by moving beyond traditional human interpretation toward data-driven, intelligent systems that support clinical decision-making. These technologies are increasingly being integrated into medical diagnostics and patient management systems. For instance, AI has been widely adopted in medical imaging to enhance the analysis of computed tomography (CT) scans, magnetic resonance imaging (MRI), and X-rays [1]-[3], while ML-powered wearable devices have demonstrated promising applications in continuous health monitoring and chronic disease management [4]-[6].

Pediatric asthma remains a condition that requires accurate and timely risk assessment. Diagnosing asthma and evaluating its severity in children under five years of age presents significant clinical challenges because conventional pulmonary function tests require active patient participation and precise forced expiratory maneuvers that young children cannot reliably perform. Moreover, current diagnostic tools, such as the spirometers and peak flow meters, are not only limited in availability but are also challenging to use with children under five years of age [7]. To overcome the limitations of effort-dependent diagnostic methods,

continuous physiological monitoring offers a non-invasive, objective alternative.

During acute pediatric asthma exacerbation, airway hyperresponsiveness contributes to airway inflammation, mucosal edema, bronchospasm, and mucus accumulation. This narrowing of the lower airways impedes pulmonary gas exchange, leading to ventilation-perfusion mismatch and acute hypoxemia, which can cause a reduction in peripheral blood oxygen saturation (SpO₂). Concurrently, the resulting hypoxemia and increased work of breathing activate arterial chemoreceptors, triggering a sympathetic nervous system response that induces a compensatory increase in heart rate. Integrating these physiological vital signs into low-cost, continuous Edge AI monitoring devices may enable early risk stratification and automated clinical triage at the point of care.

A recent study [8] emphasized the urgent need for effective digital health interventions that are affordable, easy to use, and capable of providing real-time insights into the severity of asthma symptoms. Despite the availability of promising data-driven techniques, a critical gap persists in translating these capabilities into deployable solutions, particularly on constrained hardware.

From a methodological standpoint, ML, particularly supervised learning, offers a variety of algorithms that can perform classification tasks with minimal human intervention [9], [10]. These algorithms include support vector machines, decision trees, K-nearest neighbors, and neural networks

[11], [12]. Among them, Gaussian Naïve Bayes (GNB) stands out due to its simplicity, speed, and competitive performance across various domains.

For example, Li and Iskander [13] compared multiple classifiers, including (GNB), for classifying sand based on particle size and shape. Although neural networks achieved the highest accuracy, they posed deployment challenges compared with simpler models such as GNB. Similarly, in [14], a Naïve Bayes-based system was developed to predict student performance with high accuracy, although its generalizability was limited by the dataset size. Other studies have applied Naïve Bayes to tasks such as optical character recognition [15] and sports analytics [16], consistently demonstrating robust accuracy, although these studies used small and domain-specific datasets.

In the medical domain, Bose et al. [17] employed Naïve Bayes to predict the long-term persistence of asthma in children using a large dataset of nearly 10,000 records. Their model provided actionable insights for early intervention, although it did not address deployment challenge. Another study [18] used ML models, including Naïve Bayes, to predict the functionality of infant incubators. Although decision trees outperformed the other classifiers, the findings nevertheless highlighted the potential of lightweight algorithms in healthcare applications.

In recent years, the emergence of Edge AI enabled trained ML models to be deployed directly onto devices, allowing real-time inference at the edge without relying on cloud-based computation [19], [20]. This development has enabled embedded devices to perform intelligent and autonomous data processing. However, it has also introduced new challenges associated with the limited RAM, processing capacity, and flash memory of embedded systems [21], [22]. In response to these challenges, this paper addresses the identified gaps by:

i) Classifying pediatric asthma risks using an efficient GNB model using heart rate and peripheral blood oxygen saturation (SpO₂) as input features.

ii) Utilizing synthetic data-generation techniques to approximate the characteristics of real-world data and improve model generalization.

GNB was chosen for its high interpretability, computational efficiency, and suitability for resource-constrained environments, making it appropriate for real-time applications in medical devices. Furthermore, the proposed system integrates synthetic data generation to address the limitation of small datasets, which is a common issue in medical research.

II. METHODOLOGY

The methodology employed in this study is presented through the structured flowchart shown in Figure 1. The method includes feature selection, data generation based on clinical guidelines, and the division of the dataset into training and testing sets. Class-specific statistical parameters are estimated during model training, after which standard performance metrics are used for evaluation. To evaluate the classifier's predictive validity and practical relevance in a healthcare setting, the results are subsequently compared with clinical benchmarks. In summary, the research involves the development of an ML model for pediatric asthma risk classification using a synthetic dataset, followed by performance evaluation.

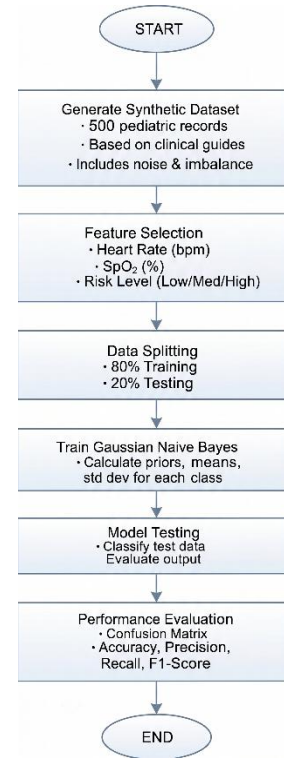


Figure 1. Flow chart of the research method

A. Dataset

A synthetic dataset comprising 500 records asthma patients aged 3-5 years was generated strictly for research purposes. The dataset was generated using GPT-4 Turbo variant through a comprehensive prompt based on the clinical guideline thresholds reported in [23] and [24]. According to [25], a dataset comprising 500 records can be sufficient for certain machine learning applications. The use of synthetic data has shown promise [26], [27] and has become a potential approach for addressing data privacy concerns [28]-[30] therefore, it was employed in this experiment to support model training. Noise and class imbalanced were also incorporated into the synthetic dataset to represent the variability commonly observed in real-world. Table 1 presents the features included in the dataset:

Table 1
Features of the dataset

Description	Features	Type
Blood oxygen (%)	SpO ₂	Numeric
Heart rate in beats per minute (bpm)	Heart_Rate	Numeric
Asthma risk level as target (Low, Moderate, High)	Risk_Level	Categorical

B. Machine Learning Model

A manual split of the dataset with an 80:20 ratio of training to testing set was used to support model generalization and reduce the risk of overfitting [31]. The training set, comprising 400 samples, was used to train the GNB classifier and develop the asthma risk classification model. The testing set, comprising 100 samples, was then used to evaluate the predictive performance of the trained model. To accelerate

training and testing, Python code was used to facilitate the model training and testing processes. The Gaussian likelihood and posterior probability were calculated using (1) and (2) respectively, to verify the classification results obtained from the training and testing dataset.

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \quad (1)$$

where:

x = feature value,

μ = mean of the feature values for a given class

σ = standard deviation of the feature values for that class,

π = constant (approximately 3.14159),

$\exp$ = is the exponential function of the natural logarithm

$$P(\text{class}|\text{feature}) = \frac{P(\text{class}) \cdot P(\text{feature}|\text{class})}{P(\text{features})} \quad (2)$$

where:

$P(\text{class}|\text{features})$ = posterior probability of class given the features,

$P(\text{class})$ = prior probability of class,

$P(\text{feature}|\text{class})$ = the likelihood which is assumed to follow a Gaussian distribution, and

$P(\text{features})$ – marginal probability of the observed features.

C. Performance evaluation metrics

To validate the reliability and effectiveness of the predictive system, several standard evaluation metrics were used. In this research, a confusion matrix, accuracy, precision, recall and F1-score were used to evaluate the model's performance. The confusion matrix summarizes the predictions produced by the Naïve Bayes classification model. An example of a confusion matrix is presented in Figure 2 [32]. For each class, the confusion matrix contains the numbers of true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN). Accuracy represents the proportion of correctly predicted observations in the test set and is calculated using (3). Precision, recall and F1-score are calculated using (4), (5), and (6), respectively.

	A	B	C
A	TP _A	FN _A	FN _A
B	FP _A	TN _A	TN _A
C	FP _A	TN _A	TN _A

Figure 2. Example of a confusion matrix

$$\text{Accuracy} = \frac{\text{Total correct prediction}}{\text{Total prediction}} \quad (3)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (5)$$

$$F1 - \text{score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

D. Data Modeling and Target Class Alignment

In addition to evaluating model accuracy, a manual comparison between the predicted risk levels and the childhood asthma severity assessment used in Malaysia clinical practice [23] was conducted to strengthen the validity of the findings. According to [33], [34] SpO2 is commonly used as an important clinical indicator of asthma severity; therefore, only SpO2 was considered in this comparison. Table 2 shows the severity levels of acute asthma exacerbations extracted from the clinical guideline.

Table 2
Severity of acute asthma exacerbations

Parameter	Mild	Moderate	Severe	Life Threatening
SpO2	>95%	92-95%	<92%	Cyanosis & <92%

To resolve discrepancies between the multiple severity categories in the clinical guidelines and the requirements of automated triage, the four clinical severity categories defined in the standard CPG guidelines were consolidated into the three-level machine learning target classification system presented in Table 3.

Table 3
ML Target Class Alignment

Clinical severity	Clinical criteria	Mapped ML Target Class
Mild	>95%	Low risk
Moderate	92-95%	Moderate risk
Severe	<92%	High Risk
Life-Threatening	Cyanosis & <92%	High Risk

E. Benchmarking Framework and Model Selection

To evaluate the classification performance and computational feasibility of the proposed Gaussian Naive Bayes (GNB) algorithm, three benchmark classifiers, namely the Support Vector Machine (SVM), Decision Tree (DT), and Random Forest (RF) were implemented under identical training and testing configurations. All models were evaluated using accuracy, precision, recall, and F1-score to assess GNB's trade-off between predictive performance and resource efficiency for deployment on microcontrollers.

III. RESULTS AND DISCUSSION

A. Proposed GNB Model Performance

The GNB model was successfully trained and evaluated using Python in JupyterLab. The training results are presented as a confusion matrix heatmap in Figure 3, while the corresponding classification report is shown in Figure 4.

Based on the confusion matrix and classification report, the precision values for the high-, low- and moderate-risk asthma classifications were 0.92, 0.88 and 0.79, respectively. The lower precision value for the moderate-risk class may be attributed to overlapping feature values and the limited distinction between this class and the high- and low-risk classes. The recall values recorded for the high-, low-, and moderate-risk classes were 0.88, 0.92 and 0.75 respectively. The lower recall value for moderate-risk class indicates that the model correctly identified only 79% of the actual moderate-risk cases. The model also achieved F1 scores of

0.90 for both the high- and the low-risk classes, and 0.77 for the moderate-risk class. These results indicate that the model was less effective in classifying moderate-risk cases but performed well in classifying high- and low-risk cases. The model achieved an overall training accuracy of 86%, indicating that 86% of the risk-level observations in the training set were correctly classified. This result demonstrates good classification performance on the training set [35].

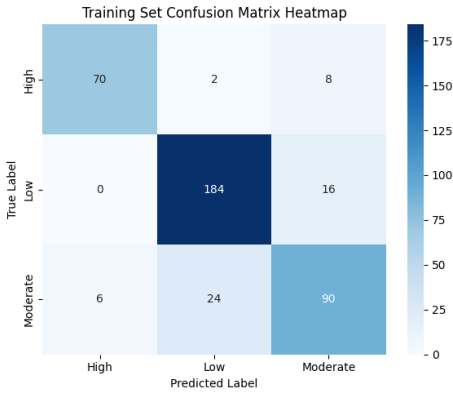


Figure 3. Confusion matrix heatmap for the training set

Classification Report:

	precision	recall	f1-score	support
High	0.92	0.88	0.90	80
Low	0.88	0.92	0.90	200
Moderate	0.79	0.75	0.77	120
accuracy			0.86	400
macro avg	0.86	0.85	0.85	400
weighted avg	0.86	0.86	0.86	400

Figure 4. Classification report for the training set

The results obtained from the test set differed from those of the training set. Based on the testing-set heatmap in Figure 5 and the classification report in Figure 6, the precision values for the high-, low- and moderate-risks classes were 0.88, 0.92 and 0.75, respectively. In terms of recall, the high-risk class recorded the lowest value of 0.70. This result is a concern in medical risk-classification context because 30% of the actual high-risk cases were not correctly identified as high risk. However, the small number of samples in this class may have contributed to fluctuations in its performance. The high F1-score for the low-risk class indicates that the model effectively classified low-risk cases. Overall, the model achieved a relatively high-test accuracy across the three classes.

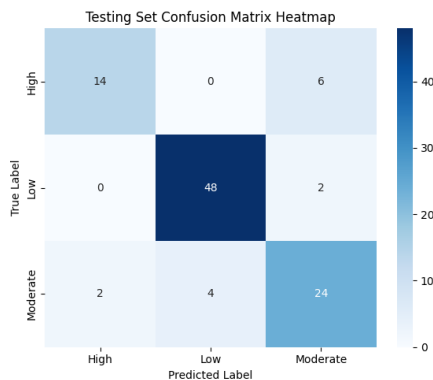


Figure 5. Confusion matrix heatmap for the test set

Classification Report:

	precision	recall	f1-score	support
High	0.88	0.70	0.78	20
Low	0.92	0.96	0.94	50
Moderate	0.75	0.80	0.77	30
accuracy			0.86	100
macro avg	0.85	0.82	0.83	100
weighted avg	0.86	0.86	0.86	100

Figure 6. Classification report for the test set

Further analysis was conducted in python to evaluate the model's behavior across different training-set sizes. The resulting learning curve, which compares the training and testing accuracies, is shown in Figure 7. The graph indicates that the model did not exhibit substantial overfitting or underfitting. Although minor fluctuations occurred at smaller sample sizes (< 150), the testing accuracy stabilized between 83.00% and 84.00%. The training accuracy stabilized at 86.00% when the full training set was used. Beyond N = 160, the difference between the training and testing accuracies remained below 0.03. This small performance gap suggests that the GNB classifier achieved a reasonable fit, with no clear evidence of substantial underfitting or overfitting.

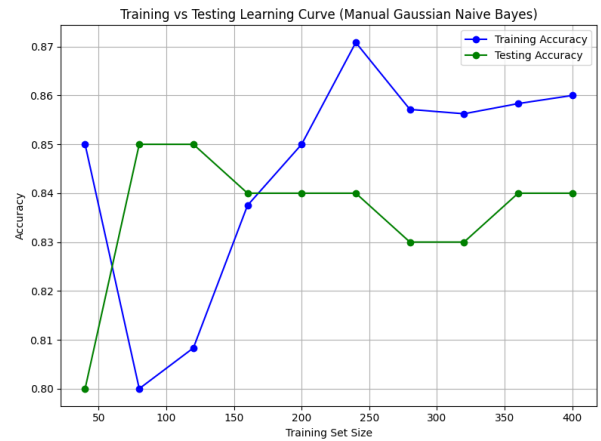


Figure 7. Learning curves for training and testing accuracy

B. Model Classification Performance Benchmark

The GNB model was also benchmarked against other classification models. As shown in Table 4, the proposed GNB model achieved the highest accuracy of 86.00%, outperforming SVM (85.00%), DT (81.00%), and RF (81.00%). GNB also recorded the highest precision and recall, although its performance advantage over SVM was limited to one percentage point.

Table 4
Model classification comparison

Classifier Model	Accuracy (%)	Precision	Recall
GNB	86.00%	0.86	0.86
SVM	85.00%	0.85	0.85
DT	81.00%	0.81	0.81
RF	81.00%	0.81	0.81

C. Predicted risk and clinical practice guideline

Examples of the predicted risk levels for 10 samples from the test dataset are shown in Table 5. A comparison between the predicted risk levels and the severity categories derived

from the clinical practice guidelines (CPG) is provided in Table 6. Based on the assessment of acute asthma exacerbation severity in children described in [23], SpO₂ was used as the clinical indicator for the comparison. For the 100 test samples, the predicted risk levels achieved 76% agreement rate with the risk categories derived from the CPG thresholds. This value represents agreement with the SpO₂-based CPG categories rather than external clinical accuracy.

Based on the training and testing results obtained using the GNB classifier, the accuracies were relatively consistent across both datasets. This consistency suggests that the model did not substantially overfit the synthetic training data and generalized reasonably well to the held-out synthetic test set. However, precision, recall, and F1-score show that accuracy alone may provide an incomplete assessment of class-specific performance, particularly for the high-risk class.

Table 5
Predicted asthma risk levels

No	Heart Rate	SpO ₂	Predicted Risk Level
1	81.82738011	94.43474741	('Low', {'Low': 0.0007207175114089116, 'Moderate': 0.0003013626208376094, 'High': 4.940629805878497e-06})
2	80.01173312	95.8393162	('Low', {'Low': 0.0020686395562453217, 'Moderate': 0.00012964412427038846, 'High': 1.3206996810781254e-06})
3	140.2840428	89.66495708	('High', {'Low': 3.69141594111822e-12, 'Moderate': 4.6698592176927134e-06, 'High': 0.0005951018748877887})
4	102.883581	98.86912779	('Low', {'Low': 0.001326756173327727, 'Moderate': 2.1004284540438135e-05, 'High': 6.679286069071785e-07})
5	71.13436562	94.06088754	('Low', {'Low': 0.00014328577192707496, 'Moderate': 5.341065307312459e-05, 'High': 1.1130111645653081e-06})
6	104.0727482	96.28187707	('Low', {'Low': 0.0021434655463780693, 'Moderate': 0.0007481471945172234, 'High': 1.4600441232351163e-05})
7	122.2664945	97.43641386	('Moderate', {'Low': 0.0001248675149397102, 'Moderate': 0.00015472167658738697, 'High': 1.2434721072865285e-05})
8	97.76635128	97.91587862	('Low', {'Low': 0.0032801198885599623, 'Moderate': 8.272403564670988e-05, 'High': 1.4179742549146603e-06})
9	97.62206162	97.27725419	('Low', {'Low': 0.003841165933264604, 'Moderate': 0.00019739301025881918, 'High': 2.926042695431997e-06})
10	81.82738011	94.43474741	('Low', {'Low': 0.0007207175114089116, 'Moderate': 0.0003013626208376094, 'High': 4.940629805878497e-06})

Table 6
Comparison between the predicted risk levels and severity risk from CPG

No	SpO ₂	Predicted Risk Level (Trained model)	Severity risk (Clinical practice guideline)
1	94.43474741	('Low', {'Low': 0.0007207175114089116, 'Moderate': 0.0003013626208376094, 'High': 4.940629805878497e-06})	Low
2	95.8393162	('Low', {'Low': 0.0020686395562453217, 'Moderate': 0.00012964412427038846, 'High': 1.3206996810781254e-06})	Low
3	89.66495708	('High', {'Low': 3.69141594111822e-12, 'Moderate': 4.6698592176927134e-06, 'High': 0.0005951018748877887})	High
4	98.86912779	('Low', {'Low': 0.001326756173327727, 'Moderate': 2.1004284540438135e-05, 'High': 6.679286069071785e-07})	Low
5	94.06088754	('Low', {'Low': 0.00014328577192707496, 'Moderate': 5.341065307312459e-05, 'High': 1.1130111645653081e-06})	Moderate
6	96.28187707	('Low', {'Low': 0.0021434655463780693, 'Moderate': 0.0007481471945172234, 'High': 1.4600441232351163e-05})	Low
7	97.43641386	('Moderate', {'Low': 0.0001248675149397102, 'Moderate': 0.00015472167658738697, 'High': 1.2434721072865285e-05})	Low
8	97.91587862	('Low', {'Low': 0.0032801198885599623, 'Moderate': 8.272403564670988e-05, 'High': 1.4179742549146603e-06})	Low
9	97.27725419	('Low', {'Low': 0.003841165933264604, 'Moderate': 0.00019739301025881918, 'High': 2.926042695431997e-06})	Low
10	94.43474741	('Low', {'Low': 0.0007207175114089116, 'Moderate': 0.0003013626208376094, 'High': 4.940629805878497e-06})	Low

Conversely, the model performed well in classifying the low-risk class. Maximizing recall, or sensitivity, for the high-risk class is clinically important because misclassifying a severe exacerbation as moderate or low risk produces false-negative outcomes in an asthma triage context [36]. To reduce this risk in point-of-care Edge AI applications, the decision threshold for the high-risk class could be lowered [37], [38]. For example, a threshold for $p \geq 0.50$ to $p \geq 0.30$ could be considered, but its exact value should be selected and validated using an independent validation dataset to balance sensitivity and false-alert rates.

Additionally, the model could be improved by providing better representation of the high-risk class in the datasets. Greater emphasis should also be placed on the high-risk class [39]–[42] because it represents potentially life-threatening asthma cases. For GNB, this can be achieved by adjusting the class prior probabilities or applying appropriate resampling techniques. The contribution of SpO₂ should be evaluated separately because it is an important clinical indicator for asthma severity, rather than treated as a class weight. Other models, such as Random Forest and Support Vector Machine (SVM) may improve classification performance; however, their deployment on embedded systems may be more complex than that of GNB.

CONCLUSION

This study demonstrates that the GNB classifier can successfully classify pediatric asthma risk using synthetic data, achieving an accuracy of 86%. Compared with the risk categories derived from the primary clinical indicator, SpO₂, the trained model achieved a 76% agreement rate.

For future work, the asthma risk predicted algorithm can be deployed on embedded hardware to develop a reliable Edge-AI medical device. Such a device could enable users to monitor asthma-related physiological indicators more conveniently and cost-effectively. However, it should function as a decision-support tool rather than replace assessment by healthcare professionals.

Although the use of synthetic data is emerging and promising, future research may incorporate Internet of Things (IoT) monitoring to collect real-world data for model validation. A wider range of appropriately selected and clinically relevant features should also be considered. The integration of AI, ML and IoT will contribute to more reliable asthma risk detection in children. Further, comparisons among different classification models could also be conducted to identify the most effective model for predicting paediatric asthma risk levels.

CONFLICT OF INTEREST

Authors declare that there is no conflict of interest regarding the publication of the paper.

AUTHOR CONTRIBUTION

The authors confirm contribution to the paper as follows: study conception and design: Dyg Khayrunsalihaty Bariyyah binti Abang Othman; data training: Dyg Khayrunsalihaty Bariyyah binti Abang Othman; analysis and interpretation of findings: Dyg Khayrunsalihaty Bariyyah binti Abang Othman, Azarina binti Azman; draft manuscript preparation:

Azarina binti Azman. All authors reviewed the findings and approved of the final manuscript.

REFERENCES

- [1] J. Iqbal et al., "Reimagining healthcare: unleashing the power of artificial intelligence in medicine," *Cureus*, vol. 15, no. 9, 2023, <https://doi.org/10.7759/cureus.44658>.
- [2] T. Habuza et al., "AI applications in robotics, diagnostic image analysis and precision medicine: current limitations, future trends, guidelines on CAD systems for medicine," *Informatics in Medicine Unlocked*, vol. 24, 2021, <https://doi.org/10.1016/j.imu.2021.100596>.
- [3] Kaplan et al., "Artificial intelligence/machine learning in respiratory medicine and potential role in asthma and COPD diagnosis," *Journal of Allergy and Clinical Immunology: In Practice*, vol. 9, no. 6, pp. 2255–2261, 2021, <https://doi.org/10.1016/j.jaip.2021.02.014>.
- [4] K. Stokes et al., "The use of artificial intelligence systems in diagnosis of pneumonia via signs and symptoms: a systematic review," *Biomed Signal Process Control*, vol. 72, 2022, <https://doi.org/10.1016/j.bspc.2021.103325>.
- [5] Z. Ashfaq et al., "A review of enabling technologies for internet of medical things (IoMT) ecosystem," *Ain Shams Engineering Journal*, vol. 13, no. 4, 2022, <https://doi.org/10.1016/j.asej.2021.101660>.
- [6] B. Paneru, B. Paneru, S. C. Sapkota, and R. Poudyal, "Enhancing healthcare with AI: sustainable AI and IoT-powered ecosystem for patient aid and interpretability analysis using SHAP," *Measurement: Sensors*, vol. 36, 2024, <https://doi.org/10.1016/j.measen.2024.101305>.
- [7] J. K. Jose, and A. Namboodiripad, "Validation of the paediatric asthma score (PAS) in evaluation of acute exacerbation of asthma in children," *Journal of Nepal Paediatric Society*, vol. 43, no. 1, pp. 95–98, 2023, <https://doi.org/10.60086/jnps495>.
- [8] G. Mosnaim et al., "Digital health technology in asthma: a comprehensive coping review," *Journal of Allergy and Clinical Immunology: In Practice*, vol. 9, no. 6, pp. 2377–2398, 2021, <https://doi.org/10.1016/j.jaip.2021.02.028>.
- [9] A. F. A. H. Alnuaimi, and T. H. K. Albaldawi, "An overview of machine learning classification techniques," *BIO Web of Conferences*, 2024, <https://doi.org/10.1051/bioconf/20249700133>.
- [10] N. H. Ali, M. E. Abdulmunem, and A. E. Ali, "Learning evolution: a survey," *Iraqi Journal of Science*, vol. 62, no. 12, pp. 4978–4987, 2021, <https://doi.org/10.24996/ijis.2021.62.12.34>.
- [11] Scikit Learn, "scikit-learn machine learning in python." Accessed: May 19, 2025, [Online]. Available: https://scikit-learn.org/stable/modules/naive_bayes.html#gaussian-naive-bayes.
- [12] F. Pedregosa et al., "Scikit-learn: machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2839, 2011.
- [13] L. Li, and M. Iskander, "Use of machine learning for classification of sand particles," *Acta Geotech*, vol. 17, no. 10, pp. 4739–4759, 2022, <https://doi.org/10.1007/s11440-021-01443-y>.
- [14] V. Nakhipova et al., "Use of the naive bayes classifier algorithm in machine learning for student performance prediction," *International Journal of Information and Education Technology*, vol. 14, no. 1, pp. 92–98, 2024, <https://doi.org/10.18178/ijiet.2024.14.1.2028>.
- [15] Hubert, P. Phoenix, R. Sudaryono, and D. Suhartono, "Classifying promotion images using optical character recognition and naïve bayes classifier," *Procedia Computer Science*, pp. 498–506, 2021, <https://doi.org/10.1016/j.procs.2021.01.033>.
- [16] A. Venkatachalam, R. D., and S. N. S. S., "Football prediction system using gaussian naïve bayes algorithm," in *Proc. of 2023 2nd International Conference on Electronics and Renewable Systems (ICEARS)*, 2023, pp. 1640–1643, <https://doi.org/10.1109/ICEARS56392.2023.10085510>.
- [17] S. Bose, C. C. Kenyon, and A. J. Masino, "Personalized prediction of early childhood asthma persistence: a machine learning approach," *PLoS ONE*, vol. 16, no. 3, 2021, <https://doi.org/10.1371/journal.pone.0247784>.
- [18] Ž. Kovačević, L. Gurbeta Pokvić, L. Spahić, and A. Badnjević, "Prediction of medical device performance using machine learning techniques: infant incubator case study," *Health and Technology*, vol. 10, pp. 151–155, 2020, <https://doi.org/10.1007/s12553-019-00386-5>.
- [19] R. Singh and S. S. Gill, "Edge AI: a survey," *Internet of Things and Cyber-Physical Systems*, vol. 3, pp. 71–92, 2023, <https://doi.org/10.1016/j.iotcps.2023.02.004>.
- [20] K. B. Letaief, Y. Shi, J. Lu, and J. Lu, "Edge artificial intelligence for 6G: vision, enabling technologies, and applications," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 1, pp. 5–36, 2022, <https://doi.org/10.1109/JSAC.2021.3126076>.

- [21] A. Paleyes, R. G. Urma, and N. D. Lawrence, "Challenges in deploying machine learning: a survey of case studies," *ACM Computing Surveys*, vol. 55, no. 6, 2022, <https://doi.org/10.1145/3533378>
- [22] S. S. Saha, S. S. Sandha, and M. Srivastava, "Machine learning for microcontroller-class hardware: a review," *IEEE Sensors Journal*, vol. 22, no. 22, pp. 21362–21390, 2022, <https://doi.org/10.1109/JSEN.2022.3210773>.
- [23] A. Omar, N. M. Zainudin, A. F. Abdullah, K. C. Tuan, P. Chan, and R. A. Awang, "Clinical Practice Guidelines for the Management of Childhood Asthma," Accessed: May 19, 2025, [Online.] Available: <https://mpaeds.my/clinical-practice-guidelines-for-the-management-of-childhood-asthma/>.
- [24] Elsevier Clinical Skills, "Emergency Secondary Assessment (Pediatric)-CE," Accessed: May 19, 2025, [Online.] Available: https://repository.phb.ac.id/740/1/Emergency-Secondary-Assessment-Skill-Pediatric-COVID-19-toolkit_010320.pdf.
- [25] K. Zantvoort, B. Nacke, D. Görlich, S. Hornstein, C. Jacobi, and B. Funk, "Estimation of minimal data sets sizes for machine learning predictions in digital mental health interventions," *npj Digital Medicine*, vol. 7, 2024, <https://doi.org/10.1038/s41746-024-01360-w>.
- [26] R. Lingo, "Exploring the potential of AI-generated synthetic datasets: a case study on telematics data with ChatGPT," *arXiv*, 2023, <https://doi.org/10.48550/arXiv.2306.13700>.
- [27] H. Hassani, and E. S. Silva, "The role of ChatGPT in data science: how AI-assisted conversational interfaces are revolutionizing the field," *Big Data and Cognitive Computing*, vol. 7, no. 2, 2023, <https://doi.org/10.3390/bdcc7020062>.
- [28] A. S. Dina, A. B. Siddique, and D. Manivannan, "Effect of balancing data using synthetic data on the performance of machine learning classifiers for intrusion detection in computer networks," *IEEE Access*, vol. 10, pp. 96731–96747, 2022, <https://doi.org/10.1109/ACCESS.2022.3205337>.
- [29] Y. Lu et al., "Machine learning for synthetic data generation: a review," *arXiv*, 2023, <https://doi.org/10.48550/arXiv.2302.04062>.
- [30] D. Rankin, M. Black, R. Bond, J. Wallace, M. Mulvenna, and G. Epelde, "Reliability of supervised machine learning using synthetic data in health care: model to preserve privacy for data sharing," *JMIR Med Inform*, vol. 8, no. 7, 2020, <https://doi.org/10.2196/18910>.
- [31] I. O. Muraina, "Ideal dataset splitting ratios in machine learning algorithms: general concerns for data analysts," in *Proc. of 7th International Mardin Artuklu Scientific Researches Conference*, 2022, pp. 496–504.
- [32] M. F. Amin, "Confusion matrix in three-class classification problems: a step-by-step tutorial," *Journal of Engineering Research*, vol. 7, 2023, <https://doi.org/10.21608/erjeng.2023.296718>.
- [33] L. Taylor, X. Ding, D. Clifton, and H. Lu, "Wearable vital signs monitoring for patients with asthma: A review," *IEEE Sensors Journal*, vol. 23, no. 3, pp. 1734–1751, 2023, <https://doi.org/10.1109/JSEN.2022.3224411>.
- [34] U. C. Onubogu, and A. Ayuk, "Pulse oximetry and peak expiratory flow rate correlations in acute asthma exacerbation in children," *Nigerian Journal of Clinical Practice*, vol. 25, no. 11, pp. 1896–1903, 2022, https://doi.org/10.4103/njcp.njcp_376_22.
- [35] J. C. Obi, "A comparative study of several classification metrics and their performances on data," *World Journal of Advanced Engineering Technology and Sciences*, vol. 8, no. 1, pp. 308–314, 2023, <https://doi.org/10.30574/wjaets.2023.8.1.0054>.
- [36] H. Jiang et al., "Machine learning-based models to support decision-making in emergency department triage," *International Journal of Medical Informatics*, vol. 145, 2021, <https://doi.org/10.1016/j.ijmedinf.2020.104326>.
- [37] C. Ganss, K. Jung, L. Schilling, S. Sonderegger, and K. W. Neuhaus, "Human and artificial intelligence performance in radiographic caries detection: ex vivo tooth section-referenced evaluation and implications for clinical decision-making," *Journal of Dentistry*, vol. 168, 2026, <https://doi.org/10.1016/j.jdent.2026.106602>.
- [38] S. Gürlür, Ö. E. Oruç, D. Göksülük, and A. Suner, "A new perspective for area under the curve in decision making with extra safety threshold value: a simulation study," *Türkiye Klinikleri Journal of Biostatistics*, vol. 14, no. 3, pp. 137–146, 2022, <https://doi.org/10.5336/biostatic.2022-89951>.
- [39] D. Xu, Y. Ye, and C. Ruan, "Understanding the role of importance weighting for deep learning," *arXiv*, 2021, <https://doi.org/10.48550/arXiv.2103.15209>.
- [40] Z. Chkirbene et al., "A weighted machine learning-based attacks classification to alleviating class imbalance," *IEEE Systems Journal*, vol. 15, no. 4, pp. 4780–4791, 2021.
- [41] R. A. Disha and S. Waheed, "Performance analysis of machine learning models for intrusion detection system using Gini impurity-based weighted random forest (GIWRF) feature selection technique," *Cybersecurity*, vol. 5, 2022, <https://doi.org/10.1186/s42400-021-00103-8>.
- [42] T. Talaei Khoei, H. Ould Slimane, and N. Kaabouch, "Deep learning: systematic review, models, challenges, and research directions," *Neural Computing and Applications*, vol. 35, pp. 23103–23124, 2023, <https://doi.org/10.1007/s00521-023-08957-4>.