Open Access Emergency Medicine

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Examining the Use of Machine Learning Algorithms to Enhance the Pediatric Triaging Approach

Hussain J Aljubran[1], Maitham J Aljubran[2], Ahmed M AlAwami[1], Mohammad J Aljubran[3], Mohammed A Alkhalifah[4], Moayd M Alkhalifah[5], Ahmed S Alkhalifah[6], Tawfik S Alabdullah[7]

[1]College of Medicine, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia; [2]Pediatric Department, King Faisal Specialist Hospital & Research Centre, Riyadh, Saudi Arabia; [3]Department of Energy Science & Engineering, Stanford University, Stanford, CA, USA; [4]Emergency Medicine Department, Johns Hopkins Aramco Healthcare, Al-Hasa, Saudi Arabia; [5]Neurology Unit, Neurosciences Institute, Johns Hopkins Aramco Healthcare, Dhahran, Saudi Arabia; [6]Pediatric Department, Qatif Central Hospital, Al Qatif, Saudi Arabia; [7]Pediatric Emergency Medicine, King Faisal Specialist Hospital & Research Centre, Riyadh, Saudi Arabia

Correspondence: Hussain J Aljubran, College of Medicine, Imam Abdulrahman Bin Faisal University, Dammam, 34212, Saudi Arabia, Tel +966540800696, Email hussainjawad1420@gmail.com; Maitham J Aljubran, Pediatric Department, King Faisal Specialist Hospital & Research Centre, Riyadh, 12212, Saudi Arabia, Tel +966544177737, Email mjmj77550@gmail.com

Purpose: Triage systems play a vital role in effectively prioritizing patients according to the seriousness of their condition. However, conventional emergency triage systems in pediatric care predominantly rely on subjective evaluations. Machine learning technologies have shown significant potential in various medical fields, including pediatric emergency medicine. Therefore, this study seeks to employ pediatric emergency department records to train machine learning algorithms and evaluate their effectiveness and outcomes in the triaging system. This model will improve accuracy in pediatric emergency triage by categorizing cases into three urgency levels (nonurgent, urgent, and emergency).

Patients and Methods: This is a retrospective observational cohort study that used emergency patient records obtained from the Emergency Department at King Faisal Specialist Hospital & Research Centre. Using the emergency severity index (a scale of 1 to 5), various machine learning techniques were employed to build different machine learning models, such as regression, instance-based, regularization, tree-based, Bayesian, dimensionality reduction, and ensemble algorithms. The accuracy of these models was compared to reach the most accurate and precise model.

Results: A total of 38,891 pediatric emergency patient records were collected. However, due to numerous outliers and incorrectly labeled data, clinical knowledge and a confident learning algorithm were employed to preprocess the dataset, leaving 18,237 patient records. Notably, ensemble algorithms surpassed other models in all evaluation metrics, with CatBoost achieving an F-1 score of 90%. Importantly, the model never misclassified an urgent patient as nonurgent or vice versa.

Conclusion: The study successfully created a machine learning model to classify pediatric emergency department patients into three urgency levels. The model, tailored to the specific needs of pediatric patients, shows promise in improving triage accuracy and patient care in pediatric emergency departments. The implication of this model in the real-life sitting will increase the accuracy of the pediatric emergency triage and will reduce the possibilities of over or under triaging.

Keywords: machine learning, artificial intelligence, emergency medicine, pediatric, triage

Introduction

Machine learning has garnered considerable attention and utilization within the healthcare sector, with the ability to enhance decision-making, improve precision in medical diagnoses, and facilitate the analysis of intricate medical data.[1–5] Machine learning algorithms have the capacity to employ routine healthcare data for predictive purposes, analyze extensive and varied collections of data, and optimize the trade-off between delay and energy in Internet of Medical

Received: 10 October 2024
Accepted: 14 January 2025
Published: 30 January 2025

Things systems.[2,6,7] These algorithms have found applications in diverse fields, including cybersecurity systems, smart cities, healthcare, e-commerce, and agriculture.[7]

In pediatric emergency departments (PED), triage systems play a vital role in effectively prioritizing patients according to the seriousness of their condition. These systems are critical for nurses and healthcare providers in PEDs, as they enable swift identification of children in need of immediate care and those who can afford to wait.[8,9] Conventional emergency triage systems in pediatric care predominantly rely on subjective evaluations and prioritize patients based on limited information. Examples of such systems include the pediatric emergency severity index (ESI), which is susceptible to the healthcare provider's judgment and exhibits suboptimal discriminatory capability.[10] Another example is the Manchester Triage System, which depends mainly on the presenting signs and symptoms of these patients, without relying on the underlying diagnosis.[11] These systems typically lack the ability to perform extensive objective data analysis and utilize predictive modeling, limiting their accuracy and efficacy in identifying patients who require immediate attention or critical care.[12]

Machine learning techniques have the ability to improve systems for prioritizing patient care in PEDs. By utilizing objective data analysis and predictive modeling, machine learning has the potential to enhance the triage methods used to determine the order in which pediatric patients receive care.[10] Machine learning models can generate more precise forecasts of medical outcomes and hospital admission requirements by taking advantage of objective data. By relying on these objective data, machine learning algorithms can guide more accurate hospital admission decision and clinical outcomes.[13] However, it is essential to acknowledge the challenges in applying machine learning algorithms to the PED triage systems, such as data imbalance, interpretability, and ethical considerations. To address these issues, this study used performance metrics, such as the F1 score and precision-recall curves, which assess model performance beyond simple accuracy and highlight effectiveness in minority classes. Additionally, these algorithms were trained using various demographic and clinical data, minimizing biases in predictions. Furthermore, patient data was anonymized to protect confidentiality, and the study adhered to ethical guidelines.

The current triaging systems lack the ability to perform extensive objective data analysis, which limits their accuracy and efficacy in triaging patients in the PEDs. To address this challenge, this study aims to create a machine learning model that can classify PED patients into three separate urgency groups (non-urgent, urgent, and emergency), which will improve accuracy in pediatric emergency triaging systems. To our knowledge, this is the first study that specifically studies pediatric triage and proposes a comprehensive three-category model that will enhance the accuracy and efficiency of the triage process.

Materials and Methods

This retrospective study did not involve any patient identification information. All patient data were anonymized and utilized solely for the research objectives. In accordance with the principles of the Declaration of Helsinki, King Faisal Specialist Hospital and Research Centre (KFSH&RC) Institutional Review Board granted approval for this study (IRB no. 2190019), ensuring the safeguarding of patient privacy.

The study dataset comprised 38,891 pediatric (age <18) emergency patient records collected from KFSH&RC. Each patient record contained administrative, demographic, and clinical information. We applied several exclusion criteria to filter our patient data. Specifically, we excluded patients who were deceased when they arrived at the hospital. We also excluded patients who were transferred to another medical facility for treatment or admitted to a psychiatric facility. Further, we did not include patients who chose to leave the hospital before their treatment was completed or who elected to leave the emergency department before being evaluated by a physician. By applying these exclusion criteria, we analyzed a cohort of patients who received complete emergency department care and treatment at the hospital facility. We further refined our patient dataset by removing records with missing or clinically implausible data values (Table 1). As a result, the total number of pediatric patient records in the dataset was reduced to 20,317.

In the proposed machine learning model, input features were derived from information commonly found in PED triage environments, such as gender, age, mode of arrival, the count of PED visits within the last 72 hours as a proxy for acuity, time of arrival, temperature, systolic blood pressure, diastolic blood pressure, respiratory rate, and pulse rate. Note that neither the oxygen saturation nor patients' visit purpose records were available in this dataset.

Table 1 Accepted Values Range to Be Included in the Study

Values	Minimum	Maximum
Pain	0	10
Temperature	33 °C	50 °C
Systolic blood pressure	0 mmHg	250 mmHg
Diastolic blood pressure	0 mmHg	150 mmHg
Respiratory rate	0/min	80/min
Pulse rate	0/min	240/min

While more granular classifications (eg, five or more urgency levels) could offer additional detail, they were not pursued due to potential challenges in clinical implementation. Increased granularity might lead to decreased interpretability and slower decision-making in high-pressure environments. Thus, the current study pursued the development of a machine learning model that categorizes PED cases into three distinct urgency categories: nonurgent, urgent, and emergency. The three-category system aligns closely with established triage protocols, such as the ESI, making it more practical for integration into existing workflows. This design improves upon previous approaches by providing a balance of accuracy, interpretability, and usability in real-world settings. Additionally, this represents a more expansive and comprehensive typology compared to the simpler two-group classifications examined in most earlier studies. The three-category model developed here aims to capture important subtleties and variations in patient acuity levels that a dichotomous model cannot adequately represent. The dataset used in the study included the ESI provided by nurses on a scale of 1 to 5. To prepare the model's output, we mapped the ESI levels to specific classifications. Nonurgent cases were categorized as ESI levels 5 or 4, urgent cases as ESI level 3, and emergency cases as ESI levels 2 or 1. Grouping ESI levels into broader categories is intended to minimize the effects of this subjective variability in triage assessments. Categorization helps ensure more consistent outcomes by reducing subtle differences in how nurses might apply the ESI scale. For simplicity, the urgency categories were called A, B, and C, where A = nonurgent (ESI 4–5), B = urgent (ESI 3), and C = emergency (ESI 1–2). To treat data imbalance between different categories, this study utilized the oversampling technique.

To validate the accuracy of the model's classifications, three independent emergency medicine consultants randomly reviewed and assessed the categorizations provided by the model. Their input confirmed the model's alignment with clinical expectations, ensuring its reliability and applicability in real-world pediatric emergency settings. Also, it is important to explore and prepare the dataset before training machine learning models. Figure 1 illustrates the relationship and distribution of the numeric input variables. Most of the plots indicating the association between the two variables exhibit a nearly normal bell-curve-shaped distribution of the three outcome categories (A, B, and C). Instances that are not urgent tend to cluster near the center. Furthermore, we examined how each input feature was distributed individually.

Whereas we filtered the features inputted to the model (eg, vital signs), to screen and detect records with incorrect labels, we use confident learning, an automated machine learning technique as it results in more robust and reliable artificial intelligence (AI) models.[14] Figure 2 demonstrates the technique's objective, which is to estimate a joint distribution between the given or noisy labels and the unknown or uncorrupted labels. This estimation is made under the assumption of a class-conditional classification noise process.[15] In applying confident learning, we used Gaussian Naive Bayes (Gaussian NB) as the backbone machine learning algorithm for its simplicity and effectiveness in handling of normally distributed and differently distributed features. Gaussian NB calculates class probabilities from feature likelihoods.[16] The purpose of using Gaussian Naive Bayes in this study was to identify mislabeled records and clean the dataset.

Based on the confident learning step, we eliminated 2080 records, reducing the dataset size to 18,237 records. The records that were eliminated were those with which the Gaussian NB model was not very confident in the triage level it

 https://doi.org/10.2147/OAEM.S494280

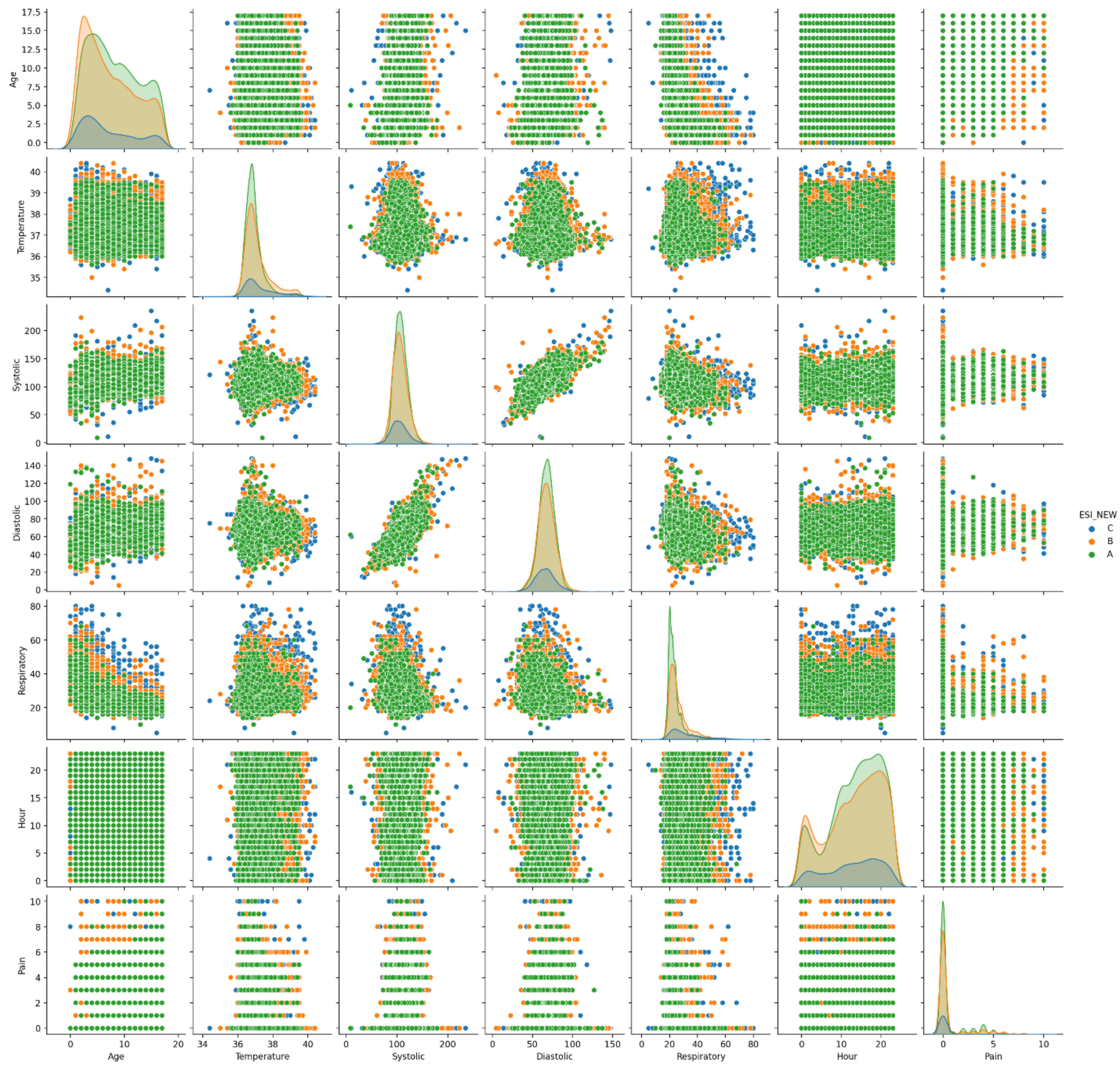

Figure 1 Illustration of the binary correlation and distribution plots of the model input features, specifically based on the raw patient records.

had predicted for them during cross-validation. Given that the model's predictions for those records were uncertain, they were filtered out. By removing the records with which the model lacked confidence in its predictions, the resulting smaller dataset included only the records where the model predicted triage levels with higher confidence. Also, this study used supervised learning where both the input and output quantities are used in training the machine learning model.

Figure 3 displays the binary correlation and distribution plots of the quantitative input features. These plots depict the relationships among the input features after data preprocessing. Our observations indicated that applying confident learning to clean the dataset resulted in improved information consistency and the elimination of outliers. Bivariate plots showed that there was an overlap, with A overlying B and B overlying C. Additionally, univariate plots demonstrated that the filtered data, both before and after applying confident learning, maintained a similar mean and variance for each of the three categories. However, confident learning tends to filter more instances from classes A and B, which have a more significant overlap. Moreover, the application of confident learning filters tended to limit categories A and B to lower magnitudes in quantities such as temperature,

 https://doi.org/10.2147/OAEM.S494280

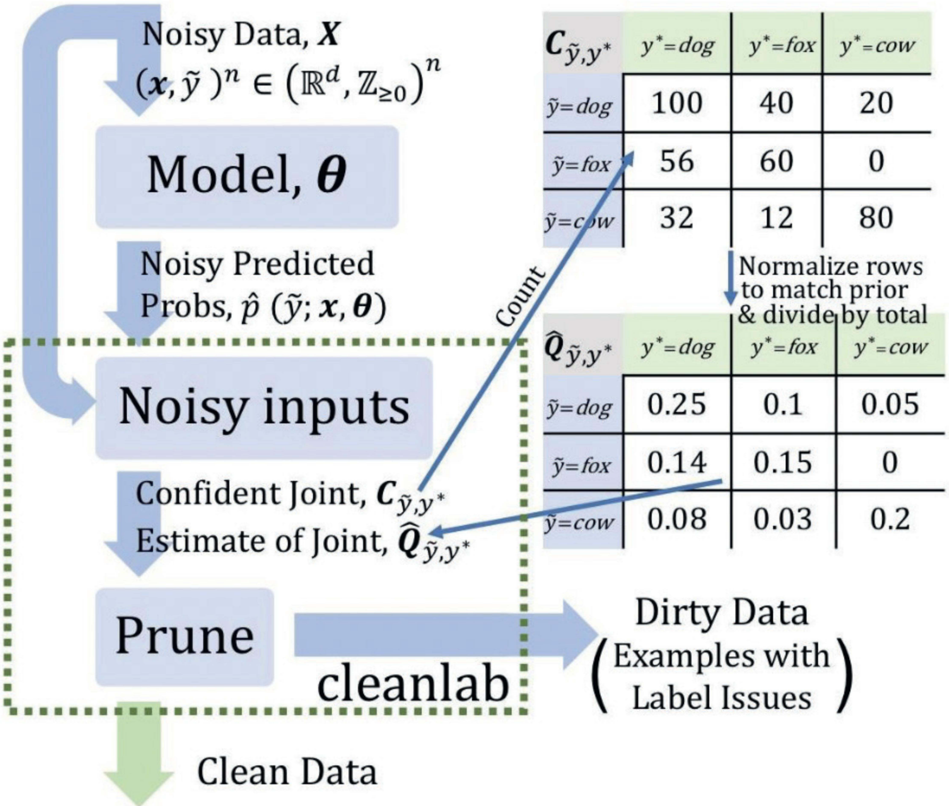

Figure 2 Confident process is applied to identify and remove mislabeled emergency triage outcomes within the study dataset.

respiratory rate, and hour of arrival. This is reasonable, as higher magnitudes of such vital signs are clinically indicative of more urgent cases under category C. In Figure 4, the dataset was cleaned using confident learning, which involved comparing the quantitative feature distributions before and after the cleaning process. The results of this comparison verify that the feature distributions remained unchanged after applying confident learning to clean the dataset.

We converted categorical data into numeric labels before building the machine learning models. This allowed the models to use those categorical features. The normalization of numeric features was carried out to enhance the efficiency of parameter optimization during training. By reducing oscillation prior to reaching the minimum loss, this normalization technique helps to reduce the time required to find the optimal parameters.[17] This process transforms the cost function into a circular shape in two dimensions and a spherical shape in three dimensions. As a result, the optimizer can converge more quickly, requiring fewer iterations to reach the optimal solution.[18] The input features were normalized and scaled using their respective mean and standard deviation statistics as computed based on the training data split. These statistics were also utilized to similarly normalize the validation and testing data splits.

We evaluated and compared the effectiveness of various types of machine learning classification algorithms for predicting triage levels of patients arriving at the PED. The algorithms employed in the study encompassed a range of techniques, such as regression (eg, logistic regression), instance-based methods (eg, k-nearest neighbors and support vector machines), regularization approaches (eg, ridge classification), tree-based models (eg, decision trees), Bayesian methods (eg, Gaussian NB), dimensionality reduction techniques (eg, linear discriminant analysis and quadratic discriminant analysis), and ensemble algorithms (eg, random forest, extra trees, boosting, Ada Boost, and gradient boosting machines). As the dataset had an unequal distribution of classes, we assessed the models using a variety of metrics to

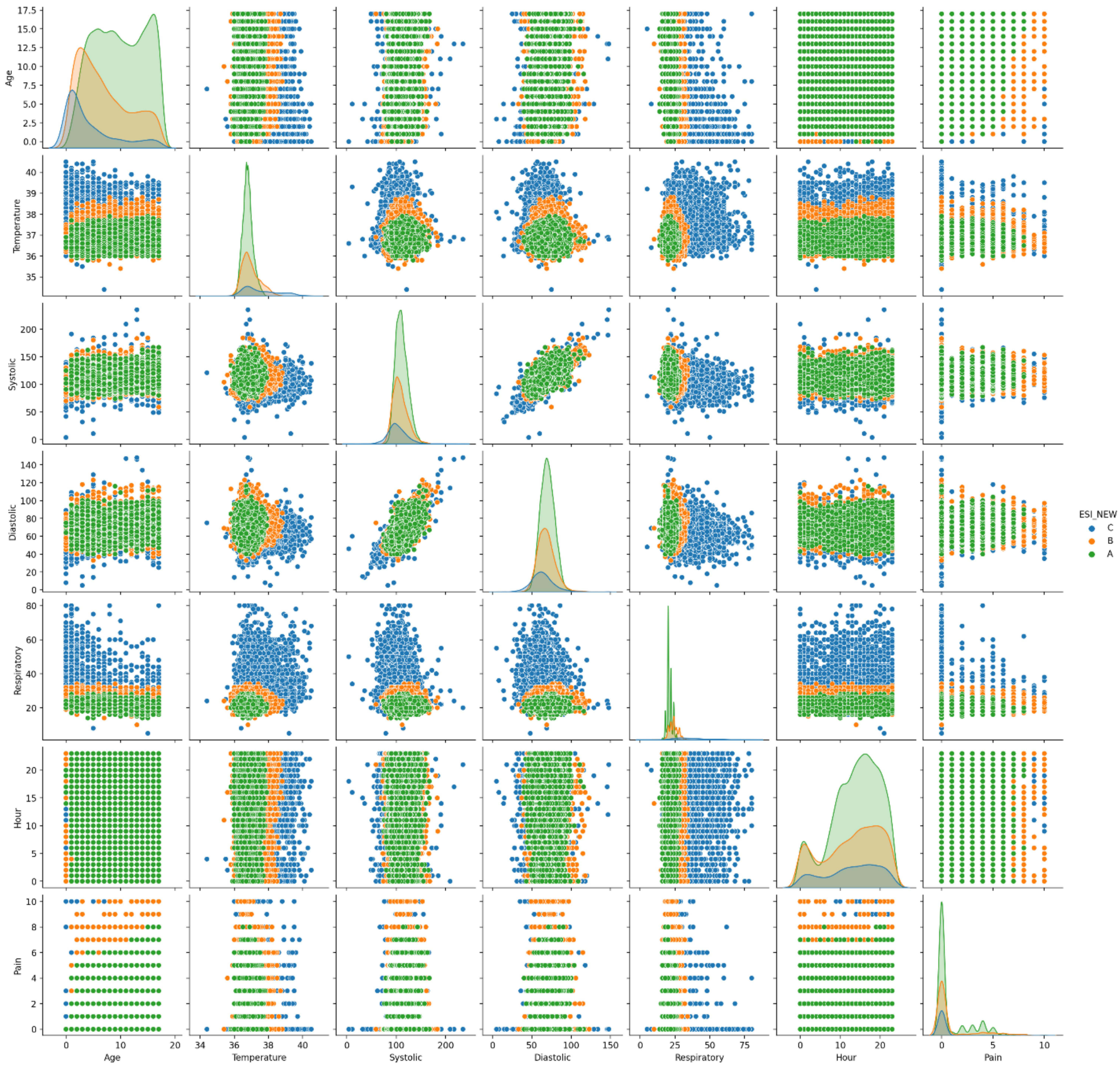

Figure 3 The binary correlation and distribution plots of the model's input characteristics based on the patient records that have been previously processed.

obtain a balanced evaluation. These metrics included accuracy, recall, precision, F-1 score, and quadratic weighted kappa.

Results

The dataset was analyzed, and for the purpose of training and testing, it was divided into two parts: 90% of the records (16,413) were assigned for training, while the remaining 10% (1824) were kept aside for unseen hold-out testing. Before building machine learning model, we examined how each input feature was distributed individually in each of the three triage outcomes (A, B, and C), and we found that for some features the probability was distinguishable. However, for other features, the outcome distributions overlapped and were indistinguishable. This indicates that properly assigning triage outcomes usually relies on using multiple input features together, not features in isolation.

To create a machine learning model with good generalizability to unseen emergency patient data, we utilized 10-fold cross-validation during the training process. 10-fold cross-validation helps ensure that the model's performance can be

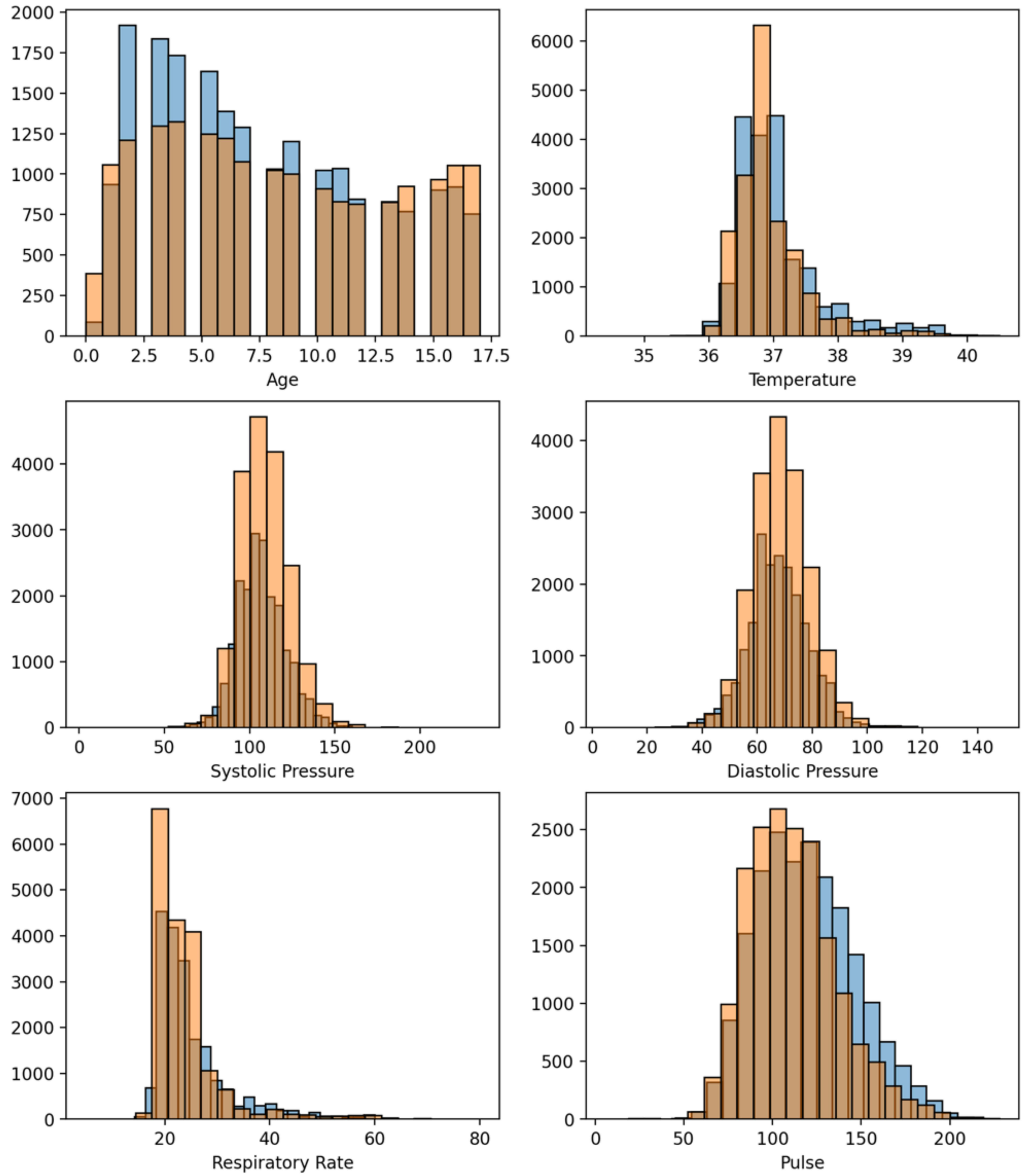

Figure 4 Side-by-side comparison of key triage prediction variables both before (in blue) and after (in Orange) the application of the confident learning algorithm.

extended to data it was not trained on, reducing overfitting. Multiple algorithms were trained, and their 10-fold cross-validation metrics were compared to determine the best-performing model for predicting outcomes for new emergency patient records. Table 2 showcases the results of our experimentation with different machine learning techniques, as each was trained using 10-fold cross-validation. The study showed that ensemble algorithms performed much better than the other techniques across all evaluation metrics, as it has been shown that CatBoost ensemble algorithm achieved the

 https://doi.org/10.2147/OAEM.S494280

Table 2 Comparison of the Effectiveness of Several Machine Learning Techniques

Model	Accuracy	AUC	Recall	Prec.	F1
CatBoost Classifier	0.9002	0.9798	0.8869	0.9008	0.9001
Light Gradient Boosting Machine	0.8969	0.9784	0.8828	0.8974	0.8966
Extreme Gradient Boosting	0.8968	0.9778	0.8814	0.8972	0.8965
Gradient Boosting Classifier	0.8901	0.9752	0.8689	0.8906	0.8896
Random Forest Classifier	0.8867	0.9735	0.864	0.8881	0.8864
Extra Trees Classifier	0.8675	0.9637	0.8294	0.8685	0.8656
Logistic Regression	0.8666	0.9534	0.8427	0.867	0.866
Decision Tree Classifier	0.8485	0.8655	0.8415	0.8485	0.8484
Ada Boost Classifier	0.8466	0.8736	0.7997	0.8456	0.8443
SVM - Linear Kernel	0.8399	0	0.8036	0.8414	0.832
Linear Discriminant Analysis	0.8259	0.9312	0.7679	0.8292	0.8224
K Neighbors Classifier	0.821	0.9147	0.7676	0.8232	0.815
Ridge Classifier	0.7671	0	0.6804	0.7736	0.75
Naive Bayes	0.6857	0.7845	0.6002	0.728	0.6132
Quadratic Discriminant Analysis	0.3336	0.5302	0.3709	0.4947	0.347

highest F1 scores. Also, we made some adjustments to the settings used to train the CatBoost model. This fine-tuning of the hyperparameters resulted in a slight gain in performance metrics. Specifically, the F1 score increased to 90% which indicates that ensemble algorithms, especially CatBoost, were extremely effective at accurately classifying the data.

Figure 5 presents the evaluation metrics for each class, while Figure 6 shows the confusion matrix. These metrics and the confusion matrix were calculated on the unseen hold-out test records. The model performed best when distinguishing between non-urgent and emergency cases in the PED. Importantly, the model never incorrectly classified an emergency case as non-urgent, meaning that it did not miss any truly critical cases. Similarly, the model did not incorrectly label any non-urgent cases as emergencies.

Discussion

Pediatric triage sets itself apart from adult triage primarily because of various factors. The foremost distinction arises from the dissimilarities in disease manifestations between pediatric and adult patients.[19] Children possess distinct physiological and developmental characteristics in comparison to adults, which can complicate and prolong the process of diagnosing and triaging certain conditions, such as appendicitis.[20] Moreover, pediatric patients who experience

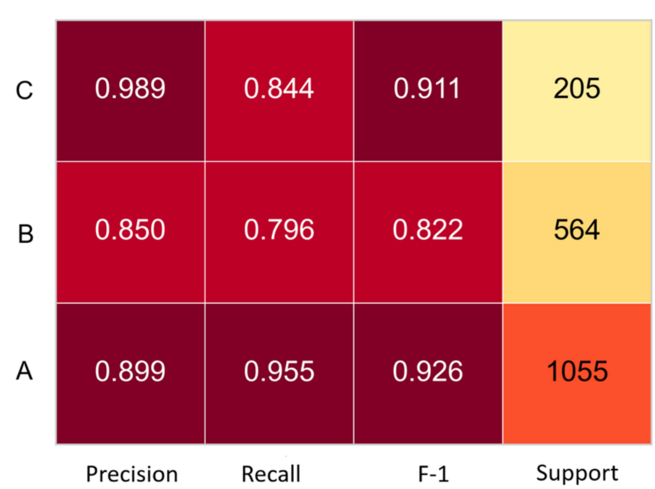

Figure 5 Model assessment metrics for the test set for each of the three emergency triage outcomes (precision, recall, and F-1 score).

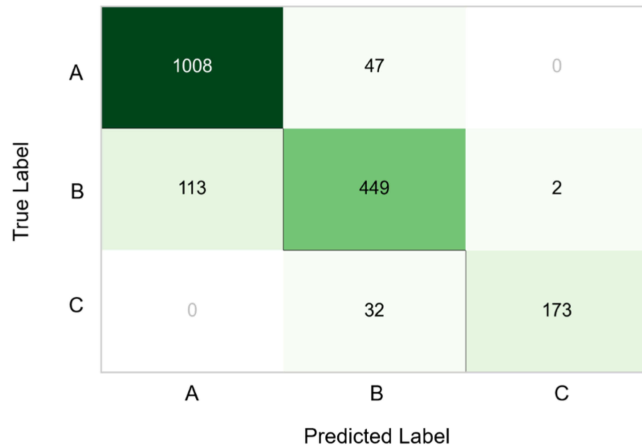

Figure 6 The confusion matrix produced while testing the trained model on an unknown test set.

traumatic injury, facing an increased risk for mortality due to their physiological differences.[21] Hence, precise and accurate triage is of paramount importance in the case of pediatric patients. The use of machine learning algorithms to improve pediatric triaging systems is a relatively new area of research. In this study, we developed a machine learning model that categorizes PED cases into three distinct urgency categories: nonurgent, urgent, and emergency.

Machine learning has emerged as a powerful tool in the medical field, revolutionizing biomedical investigations, personalized healthcare, and computer-assisted diagnosis.[22] By leveraging advanced algorithms, machine learning can integrate diverse patient data, surpassing human performance to generate accurate predictions.[23] The application of machine learning in healthcare holds the potential to enhance decision-making, refine treatment guidelines, and expedite precise diagnoses, contributing to overall healthcare advancements.[24] Within emergency medicine, machine learning has found utility in triage, risk assessment, medical imaging, and optimizing emergency department operations.[25] Notably, machine learning algorithms have been employed to anticipate health outcomes for emergency department visits, including fatalities and subsequent emergency service visits, offering valuable insights for timely intervention.[26] There are machine learning models that can forecast cardiac arrest during the conveyance of trauma patients, enabling more exact prognostication and potentially saving lives. These models may help identify high-risk patients earlier and guide time-sensitive interventions.[27]

One of the most promising applications of machine learning is its use in improving emergency department triage systems based on a huge amount of data to prioritize patients. Recent studies have focused on exploring the utility of machine learning algorithms in improving triaging systems in different populations and outcome measures. However, only Goto et al used pediatric data to develop a machine learning model to predict clinical outcomes in PEDs.[10] Although no study has focused on improving PED triaging systems in the pediatric setting, a previous study by Raita et al developed a triaging model that can improve decision-making in the adult emergency department.[28] Thus, this study makes a unique contribution by developing a comprehensive three-category triage model in the PED.

In a different context, Goto et al compared critical care outcome predictions, demonstrating machine learning models that exhibited higher discrimination ability (C statistics 0.84–0.85) compared to the reference model (C statistic of 0.78), along with higher sensitivity (0.71–0.78) compared to the reference model's sensitivity of 0.54.[10] However, the reference model had a higher specificity of 0.91, while the machine learning models had specificities of 0.77–0.86. Similarly, for the prediction of hospitalization outcomes, the machine learning models demonstrated better discrimination ability (C statistics 0.78–0.80) than the reference model (C statistic of 0.73), with sensitivities of 0.67–0.74 compared to the reference model's sensitivity of 0.83. Regarding specificity, the reference model achieved a specificity of 0.55, while the machine learning models achieved specificities ranging from 0.71 to 0.75.[10]

In a comparison of the critical care outcome with the hospitalization outcome, Raita et al showed that the machine learning model outperformed the reference model in all aspects. Both utilize the area under the curve (AUC) as a measure of their discriminatory abilities for critical care outcome; the machine learning models had higher AUC values (ranging from 0.84 to 0.86) compared to the reference model (0.74), indicating better discriminatory ability. They also exhibited higher sensitivity (ranging from 0.75 to 0.86) and lower specificity (ranging from 0.68 to 0.77) compared to the reference model (sensitivity: 0.50, specificity: 0.86). Similarly, regarding the hospitalization outcome, compared to the reference model, the machine learning models had higher AUC values (0.69 vs 0.81–0.82), lower sensitivity (0.87 vs 0.71–0.79), and higher specificity (0.42 vs 0.71–0.76).[28] These studies collectively highlight the potential of machine learning techniques in improving triage accuracy and clinical decision-making in emergency care settings.

The present study stands out for its comprehensive approach to pediatric triaging, setting it apart from Goto et al's study and Raita et al's study.[10,28] While the latter studies focused on specific outcomes or a comparison between machine learning models and conventional approaches in adult patients, this study took a different approach by developing a three-category triage model that effectively captures subtleties and variations in pediatric patient acuity levels. Moreover, this study used multiple machine learning algorithms, with CatBoost emerging as the top-performing model for predicting triage levels in the emergency department (F1 score of 90%). Hence, this innovative framework has the potential to significantly improve patient outcomes in PEDs.

Limitations

Although this study has several strengths, some limitations should be acknowledged. First, the dataset was obtained from only one center, which may limit the generalizability of the findings to other healthcare settings. Second, the study's reliance on historical data means that the model's performance may not account for potential changes in patient profiles or triage practices over time. Third, the dataset used in the study did not provide information on the clinical outcomes of the index visits for the patients.

Conclusion

This study successfully developed a machine learning model for categorizing PED patients into three urgency categories. By considering the unique needs and characteristics of pediatric patients, the model demonstrated the potential to improve triage accuracy and enhance patient care in pediatric emergency departments. While this study's findings are promising, further research and validation are necessary to ensure the model's generalizability and effectiveness in diverse pediatric emergency settings. Additionally, future studies focusing on the final patient outcome are needed to improve the direct impact of these tools in the patient outcomes.

Abbreviations

AI, artificial intelligence; AUC, area under the curve; ESI, emergency severity index; Gaussian NB, Gaussian Naive Bayes; KFSH&RC, King Faisal Specialist Hospital and Research Centre; PED, pediatric emergency department.

Disclosure

The author(s) report no conflicts of interest in this work.

References

1. Grote T, Berens P. On the ethics of algorithmic decision-making in healthcare. *J Med Ethics*. 2019;46(3):205–211. doi:10.1136/medethics-2019-105586
2. Roth JA, Battegay M, Juchler F, Vogt JE, Widmer AF. Introduction to machine learning in digital healthcare epidemiology. *Infect Control Hosp Epidemiol*. 2018;39(12):1457–1462. doi:10.1017/ice.2018.265
3. Ghaffar Nia N, Kaplanoglu E, Nasab A. Evaluation of artificial intelligence techniques in disease diagnosis and prediction. *Discov Artif Intell*. 2023;3(1):5. doi:10.1007/s44163-023-00049-5
4. Rahmani AM, Yousefpoor E, Yousefpoor MS, et al. Machine learning (ML) in medicine: review, applications, and challenges. *Mathematics*. 2021;9(22):2970. doi:10.3390/math9222970
5. Tassew T, Nie X. A comprehensive review of the application of machine learning in medicine and health care. *TechRxiv*. 2022. doi:10.36227/techrxiv.21204779.v1
6. Lakhan A, Mohammed MA, Nedoma J, et al. Federated-learning based privacy preservation and fraud-enabled blockchain IoMT system for healthcare. *IEEE J Biomed Health Inform*. 2022;27:664–672. doi:10.1109/jbhi.2022.3165945
7. Dinesh A, Rangaraj J. Energy efficient clustering and routing using hybrid fuzzy with modified rider optimization algorithm in IoT - enabled wireless body area network. *J Mach Comput*. 2023;03(02):171–183. doi:10.53759/7669/jmc202303016

8. Liu TT, Cheng CT, Hsu CP, et al. Validation of a five-level triage system in pediatric trauma and the effectiveness of triage nurse modification: a multi-center cohort analysis. *Front Med*. 2022;9:947501. doi:10.3389/fmed.2022.947501
9. Tran A, Valo P, Rouvier C, et al. Validation of the computerized pediatric triage tool, pediaTRI, in the pediatric emergency department of lenval children's hospital in nice: a cross-sectional observational study. *Front Pediatr*. 2022;10:840181. doi:10.3389/fped.2022.840181
10. Goto T, Camargo CA, Faridi MK, Freishtat RJ, Hasegawa K. Machine learning-based prediction of clinical outcomes for children during emergency department triage. *JAMA Network Open*. 2019;2(1):e186937. doi:10.1001/jamanetworkopen.2018.6937
11. van der Wulp I, van Baar ME, Schrijvers AJ. Reliability and validity of the Manchester triage system in a general emergency department patient population in the Netherlands: results of a simulation study. *Emerg Med J*. 2008;25(7):431–434. doi:10.1136/emj.2007.055228
12. Mercurio L, Pou S, Duffy S, Eickhoff C. Risk factors for pediatric sepsis in the emergency department: a machine learning pilot study. *Pediatr Emerg Care*. 2023;39(2):48–56. doi:10.1097/pec.0000000000002893
13. Patel SJ, Chamberlain DB, Chamberlain JM. A machine learning approach to predicting need for hospitalization for pediatric asthma exacerbation at the time of emergency department triage. *Acad Emerg Med*. 2018;25(12):1463–1470. doi:10.1111/acem.13655
14. Northcutt C, Jiang L, Chuang I. Confident learning: estimating uncertainty in dataset labels. *J artif intell res*. 2021;70:1373–1411. doi:10.1613/jair.1.12125
15. Angluin D, Laird P. Learning from noisy examples. *Machine Learning*. 1988;2(4):343–370. doi:10.1007/BF00116829
16. Chan TF, Golub GH, LeVeque RJ Updating formulae and a pairwise algorithm for computing sample variances. In Proceedings of the COMPSTAT 1982 5th Symposium; 30 August–3 September 1982, Toulouse, France, Cham, Switzerland:Springer;1982: pp. 30–41.
17. Grus J. *Data Science From Scratch: First Principles With Python*. Sebastopol, CA, USA: O'Reilly Media, Inc.; 2015.
18. Ristanto T, Horne R Machine learning applied to multiphase production problems. [Master's thesis]: Stanford University; Stanford, California, USA: 2018. Available from: https://pangea.stanford.edu/ERE/pdf/pereports/MS/Ristanto2018.pdf. Accessed January 21, 2025.
19. Baumann MR, Strout TD. Evaluation of the emergency severity index (version 3) triage algorithm in pediatric patients. *Acad Emerg Med*. 2005;12(3):219–224. doi:10.1197/j.aem.2004.09.023
20. Su D, Li Q, Zhang T, et al. Prediction of acute appendicitis among patients with undifferentiated abdominal pain at emergency department. *BMC Med Res Methodol*. 2022;22(1):18. doi:10.1186/s12874-021-01490-9
21. Davis M, Cayer S, Cicero M. Considerations for pediatric mass shooting triage training: a qualitative analysis. *Prehosp Disaster Med*. 2023;38(1):s215. doi:10.1002/hsr2.1629
22. Handelman GS, Kok HK, Chandra RV, Razavi AH, Lee MJ, Asadi H. eDoctor: machine learning and the future of medicine. *J Intern Med*. 2018;284(6):603–619. doi:10.1111/joim.12822
23. Fröhlich H, Balling R, Beerenwinkel N, et al. From hype to reality: data science enabling personalized medicine. *BMC Med*. 2018;16(1):150. doi:10.1186/s12916-018-1122-7
24. Koteluk O, Wartecki A, Mazurek S, Kołodziejczak I, Mackiewicz A. How do machines learn? Artificial intelligence as a new era in medicine. *J Pers Med*. 2021;11(1):32. doi:10.3390/jpm11010032
25. Mueller B, Kinoshita T, Peebles A, Graber MA, Lee S. Artificial intelligence and machine learning in emergency medicine: a narrative review. *Acute Med Surg*. 2022;9(1):740. doi:10.1002/ams2.740
26. Nunes BP, Vissoci J, Delpino FM, et al. Machine learning analysis to predict health outcomes among emergency department users in Southern Brazil: a protocol study. *Rev Bras Epidemiol*. 2021;24:210050. doi:10.1590/1980-549720210050
27. Kitano S, Ogawa K, Igarashi Y, et al. Development of a machine learning model to predict cardiac arrest during transport of trauma patients. *J Nippon Med Sch*. 2023;90(2):186–193. doi:10.1272/jnms.jnms.2023_90-206
28. Raita Y, Goto T, Faridi MK, Brown DFM, Camargo CA, Hasegawa K. Emergency department triage prediction of clinical outcomes using machine learning models. *Crit Care*. 2019;23(1):64. doi:10.1186/s13054-019-2351-7