Preprints.org

Article

Not peer-reviewed version

Comparison of the Performance of Machine Learning Models in Predicting ESI Triage Levels Using Data of Non-Traumatic Patients from the Triage Point at the Emergency Medicine Department of Lampang Hospital

Noppadon Seesuwan * , Napat Phetkub , Auswin Rojanasumapong , Nipada Chuenjai

Posted Date: 7 August 2024

doi: 10.20944/preprints202408.0428.v1

Keywords: emergency severity index; machine learning; gradient boosting; neural network; logistic regression; triage; emergency department

Article

Comparison of the Performance of Machine Learning Models in Predicting ESI Triage Levels Using Data of Non-Traumatic Patients from the Triage Point at the Emergency Medicine Department of Lampang Hospital

Noppadon Seesuwan [1,*], Napat Phetkub [2] and Auswin Rojanasumapong [2] and Nipada Chuenjai [3]

[1] Emergency Department, Lampang Hospital, Thailand 1; noppadon_kim@yahoo.com
[2] Department of Social Medicine, Lampang Hospital, Lampang, Thailand
[3] Medical Records and Medical Information Department, Lampang Hospital, Thailand
* Correspondence: noppadon_kim@yahoo.com; Tel.: +66-093-139-2567

Abstract: Emergency departments (EDs) are critical in urgent care, where accurate triage optimizes patient flow and resource allocation. Manual triage faces challenges due to increasing volume and complexity. This study compares logistic regression, gradient boosting, neural network, and random forest models in predicting Emergency Severity Index (ESI) triage levels for non-traumatic patients at Lampang Hospital's ED. Using data from January 1, 2023, to April 30, 2024, we analyzed 45,246 complete records. The gradient boosting model achieved the highest accuracy (0.81), significantly outperforming logistic regression (accuracy 0.64). Pain scale, sex, and mean arterial pressure were key predictors. This is the first comparison of these models for ESI triage in a Thai hospital, highlighting their potential to improve triage accuracy and efficiency. Implementing these models could enhance patient outcomes and resource management in ED.

Keywords: emergency severity index; machine learning; gradient boosting; neural network; logistic regression; triage; emergency department

1. Introduction

When patients come in with different degrees of medical urgency, emergency departments (EDs) are crucial for delivering the immediate care they need. Efficient triage systems are essential for prioritizing patients based on the severity of their conditions, ensuring timely and appropriate medical attention. The Emergency Severity Index (ESI) is a widely used triage tool that categorizes patients into five levels, from level 1 (most urgent) to level 5 (least urgent), based on their symptoms and resource needs [1]. Accurate ESI triage is crucial for optimizing patient flow, resource allocation, and overall ED efficiency [2].

The increasing volume of ED visits and the complexity of patient presentations pose significant challenges for manual triage, leading to potential inconsistencies and delays in patient care [3]. To tackle these issues, machine learning (ML) models have become valuable tools for accurately and consistently predicting triage levels [4]. These models can process large datasets and uncover patterns that might be missed by human triage nurses, enhancing both the reliability and efficiency of the triage process [5]. Several ML techniques have been investigated for predicting ESI triage levels, including logistic regression, decision trees, random forests, support vector machines, and neural networks. [6]. Each technique offers unique benefits and faces certain limitations, influenced by factors such as data quality, feature selection, and model complexity [7]. Comparative analyses of these models are essential for identifying the most effective approaches for specific ED settings and patient populations [8]. Recent studies have demonstrated the potential of ML models to enhance ESI triage prediction. For instance, Raita and colleagues (2019) used a gradient boosting algorithm to

 doi:10.20944/preprints202408.0428.v1

forecast ESI levels, demonstrating superior accuracy compared to traditional triage methods [8]. In A similar vein, Hong and colleagues (2020) evaluated various ML algorithms and discovered that ensemble models, which integrate the predictions of multiple base models, frequently deliver better performance in predicting triage levels [9]. Although the results are encouraging, implementing ML-based triage systems in emergency departments presents several obstacles, such as data privacy issues, the integration with current workflows, and the necessity for continuous model validation and updates [10,11]. Moreover, the variability in patient populations and ED practices necessitates the customization of ML models to ensure their applicability and effectiveness in different settings [12].

While previous studies have explored ML models for ESI triage, their application in non-Western settings, particularly in Thailand, remains underexplored. In this study, we aim to conduct a comprehensive comparative analysis of various machine learning models for predicting ESI triage levels in Lampang Hospital's emergency room settings, focusing on non-traumatic patients. By evaluating the performance of different models, we seek to identify the most accurate and efficient approaches for improving triage accuracy and enhancing ED operations.

2. Materials and Methods

This study employs a retrospective observational design to evaluate the performance of various machine learning models in predicting Emergency Severity Index (ESI) triage levels. Data were collected from the emergency department (ED) of Lampang Hospital, focusing on non-traumatic patients. The data collection period spanned from January 1, 2023, to April 30, 2024.

The dataset consists of various clinical and demographic features relevant to the triage process, including but not limited to age, sex, vital signs (blood pressure, pulse rate, respiratory rate, oxygen saturation, and temperature), Glasgow Coma Scale (GCS) score, chief complaints, pain scale, and transport method. Chief complaints were categorized using Criteria Based Dispatch (CBD) into 25 groups [13]. Each record represents an ED visit, with the corresponding ESI triage level assigned by trained triage nurses. Data preprocessing steps included handling missing values and outlier detection. Only records with complete data for all relevant features were included in the analysis. Records with any missing values were excluded to ensure the integrity and consistency of the dataset.

Criteria Based Dispatch (CBD)[10] according to the criteria for sorting and prioritizing emergency patient care as established by the Thai Emergency Medicine Foundation in 2013:

Code 1: Abdominal/Back/Pelvic and Groin Pain
Code 2: Anaphylaxis/Allergic Reactions
Code 3: Animal Bites
Code 4: Bleeding (non-traumatic)
Code 5: Breathing Difficulties
Code 6: Cardiac Arrest
Code 7: Chest Pain/Cardiac Pain
Code 8: Chocking
Code 9: Diabètes
Code 10: Environmental Hazard
Code 11: Unassigned
Code 12: Headache/Neck Pain
Code 13: Psychiatric/Behavioral Issues
Code 14: Drug Overdose/Poisoning
Code 15: Obstetric/Gynecological Emergencies
Code 16: Seizures
Code 17: General Illness/Weakness (Non-specific)/Others
Code 18: Weak Limbs/Difficulty Speaking/Facial Droop (Stroke)
Code 19: Unconscious/Unresponsive/Transient Loss of Consciousness
Code 20: Pediatric/Emergency Pediatric Care
Code 21: Assault/Injury

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2024 doi:10.20944/preprints202408.0428.v1

Code 22: Burns - Thermal/Electrical/Chemical
Code 23: Drowning/Water-Related Injuries
Code 24: Falling
Code 25: Motor Vehicle Accidents

All features in the dataset were considered for model training, including demographic information, clinical parameters, and transport details. Feature importance was assessed using various techniques, such as correlation analysis and feature importance scores from ensemble models.

Four popular machine learning models were selected for comparison based on their reported performance in previous studies and their suitability for classification tasks: Logistic Regression, Gradient Boosting, Neural Network, Random Forest

The dataset was divided into training and testing sets using stratified cross-validation with k=10 folds to ensure balanced representation of ESI levels in both sets. Model performance was evaluated using precision-recall graphs, accuracy, precision, recall, F1 score, confusion matrix, and feature importance.

Data analysis and model training were conducted using Python programming language with relevant libraries such as pandas, scikit-learn, TensorFlow, and XGBoost. Statistical analyses and data visualization were performed using seaborn and matplotlib libraries.

Ethical Considerations

This study was conducted in accordance with ethical guidelines and received approval from the institutional review board (IRB) of Lampang Hospital. Data were anonymized to protect patient privacy, and all analyses were performed on de-identified datasets.

3. Results

The initial dataset comprised 72,389 records of non-traumatic patients who visited the emergency department of Lampang Hospital. After performing data cleaning and removing records with missing values, 45,245 complete records remained for analysis.

The study population consisted of 52.6% females and 47.4% males, with an overall mean age of 52.3 years (SD ±22.2). The distribution of vital signs and other clinical parameters varied across ESI levels, with higher acuity levels (ESI-1 and ESI-2) generally showing more severe clinical presentations were shown in Table 1.

Table 1. Baseline characteristics.

Characteristics (N = 45,245)	**ESI-1 (n=4,784)**	**ESI-2 (n=25,200)**	**ESI-3 (n=10,739)**	**ESI-4 (n=3,593)**	**ESI-5 (n=929)**
Female (n=23,775)	2,148 (44.9%)	12,882 (51.1%)	6,017 (56%)	2,173 (60.5%)	555 (59.4%)
Male (n=21,470)	2,636 (55.1%)	12,318 (48.9%)	4,722 (44%)	1,420 (39.5%)	374 (40.3%)
Age (years), Mean±SD	64.6±17.8	52.3±22.2	52.2±21.3	38.5±20.2	33.5±16.5
Systolic Blood Pressure (mmHg) Mean±SD	125.9±35.9	135.1±26.6	133.2±22.4	127.5±18.7	123.8±16.3
Diastolic Blood Pressure (mmHg) Mean±SD	73.9±22.7	79.6±16.3	78.1±13.8	78.2±12.7	77.2±11.5
Mean arterial pressure (mmHg), Mean±SD	91.2±25.7	98.1±18.1	96.4±14.9	94.5±13.3	92.6±11.9
Respiratory rate, Mean±SD	26.6±7.9	20.6±3.1	19.2±1.1	19.0±1.2	18.9±1.1
Pulse rate, Mean±SD	99.8±26.7	95.3±20.9	81.9±12.4	84.0±11.8	84.3±10.7

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 August 2024 doi:10.20944/preprints202408.0428.v1

SpO2, Mean±SD	94.3±8.3	97.6±2.0	98.1±1.4	98.3±1.3	98.2±1.2
Temperature, Mean±SD	37.1±1.1	36.9±1.0	36.6±0.7	36.6±0.7	36.6±0.6
GCS Median [Min, Max]	15 [3,15]	15 [9,15]	15 [15,15]	15 [15,15]	15 [15,15]
Pain Scale Mean±SD	0 [0,0]	0 [0,5]	0 [0,5]	0 [0,4]	0 [0,2]
Shift					
Morning	2,118 (44.3%)	9,181(36.4%)	2,961 (27.6%)	752(20.9%)	266(28.6)
Evening	1,899(39.7%)	12,279(48.7%)	5,949 (55.4%)	2,222 (61.8%)	584 (62.9%)
Night	767(16.0%)	3,740(14.8%)	1,829(17.0%)	619(17.2%)	79(8.5%)
Carrier					
Refer	1,311(27.4%)	4,773(18.9%)	976(9.1%)	6(0.02%)	0(0%)
Relative	1,996(41.7%)	14,174(56.3%)	7,499(69.8%)	2,140(59.6%)	367(39.5%)
ALS	490(10.2%)	529(2.1%)	39(0.4%)	2(0.1%)	3(0.3%
BLS	686(14.3%)	2,343(9.3%)	665(6.2%)	77(2.1%)	3(0.3%)
Citizen	0(0%)	11(0.04%)	2(0.02%)	1(0.03%)	0(0%)
FR	63(1.3%)	261(1.0%)	53(0.5%)	13(0.4%)	0(0%)
Friend	29(0.6%)	508(2.0%)	210(1.9%)	113(3.2%)	29(3.1%)
Other	75(1.6%)	179(0.7%)	42(0.4%)	8(0.2%)	5(0.5%)
By yourself	134(2%)	2,422(9.6%)	1,253(11.7%)	1,233(34.3%)	522(56.2%)
Transfer					
Stretcher	4,420(92.4%)	17,882(70.9%)	6,417(59.8%)	677(18.8%)	15(1.6%)
Carry	7(0.2%)	55(0.2%)	17(0.2%)	6(0.2%)	0(0%)
Walk	126(2.6%)	4,714(18.7%)	2,715(25.3%)	2,618(72.9%)	892(96.1%)
Wheelchair	231(4.8%)	2,549(10.1%)	1,590(14.8%)	292(8.1%)	22(2.4%)
CBD					
1. Abdominal/Back/Pelvic and Groin Pain	369 (7.7%)	5878 (23.3%)	3915 (36.5%)	957 (26.6%)	60 (6.5%)
2. Anaphylaxis/Allergic Reactions	19 (0.4%)	398 (1.6%)	226 (2.1%)	101 (2.8%)	13 (1.4%)
3. Animal Bites	2 (0.04%)	10 (0.04%)	3 (0.03%)	17 (0.5%)	7 (0.8%)
4. Bleeding (non-traumatic)	96 (2%)	553 (2.2%)	245 (2.3%)	24 (0.7%)	0 (0.0%)
5. Breathing Difficulties	1972 (41.2%)	2525 (10%)	115 (1.1%)	28 (0.8%)	4 (0.4%)
6. Cardiac Arrest	51 (0.7%)	0 (0%)	0 (0%)	0 (0.0%)	0 (0%)
7. Chest Pain/Cardiac Pain	291 (6.1%)	2192 (8.7%)	275 (2.6%)	24 (0.7%)	1 (0.1%)
8. Chocking	11 (0.2%)	20 (0.1%)	10 (0.1%)	3 (0.1%)	0 (0.0%)

 doi:10.20944/preprints202408.0428.v1

9. Diabètes	17 (0.4%)	132 (0.5%)	10 (0.1%)	0 (0.0%)	1 (0.1%)
10. Environmental Hazard	3 (0.1%)	17 (0.1%)	30 (0.3%)	22 (0.6%)	7 (0.8%)
12.Headache/Neck Pain	39 (0.8%)	1071 (4.3%)	1044 (9.7%)	439 (12.2%)	75 (8.1%)
13. Psychiatric/Behavioral Issues	6 (0.1%)	291 (1.2%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
14. Drug Overdose/Poisoning	2 (0.04%)	40 (0.2%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
15. Obstetric/Gynecological Emergencies	9 (0.2%)	153 (0.6%)	80 (0.7%)	7 (0.2%)	0 (0.0%)
16. Seizures	136 (2.8%)	651 (2.6%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
17. General Illness/Weakness (Non-specific)/Others	1475 (30.8%)	8593 (34.1%)	4471 (41.6%)	1810 (50.4%)	715 (77%)
18. Weak Limbs/Difficulty Speaking/FacialDroop (Stroke)	142 (3%)	1607 (6.4%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
19.Unconscious/Unresponsive/Transient Loss of Consciousness	133 (2.8%)	357 (1.4%)	0 (0.0%)	(0.0%)	0 (0.0%)
20. Pediatric/Emergency Pediatric Care	38 (0.9%)	620 (2.7%)	207 (2.7%)	110 (4%)	34 (4.2%)

The performance of logistic regression, gradient boosting, and neural network models was compared based on precision, recall, F1 score, and accuracy. The gradient boosting model demonstrated the highest overall performance with an accuracy of 0.81, precision of 0.81, recall of 0.81, and F1 score of 0.81. The neural network model also performed well, with all metrics at 0.78. Logistic regression showed the lowest performance among the three models, with an accuracy of 0.64 and an F1 score of 0.60 were shown in Table 2.

Table 2. Comparative the overall performance of different models.

Model	Accuracy	Precision	Recall	F1 Score
Logistic Regression	0.70	0.69	0.64	0.68
Random Forest	0.80	0.80	0.80	0.80
Gradient Boosting	0.81	0.81	0.81	0.81
Neural Network	0.79	0.79	0.79	0.79

Precision-recall curves were generated for each model to evaluate their performance across different ESI levels. The gradient boosting model showed the highest precision and recall values across most ESI levels, followed by the neural network model. Logistic regression lagged, particularly in precision for lower acuity levels (ESI levels 4 and 5) were shown in Figure 1.

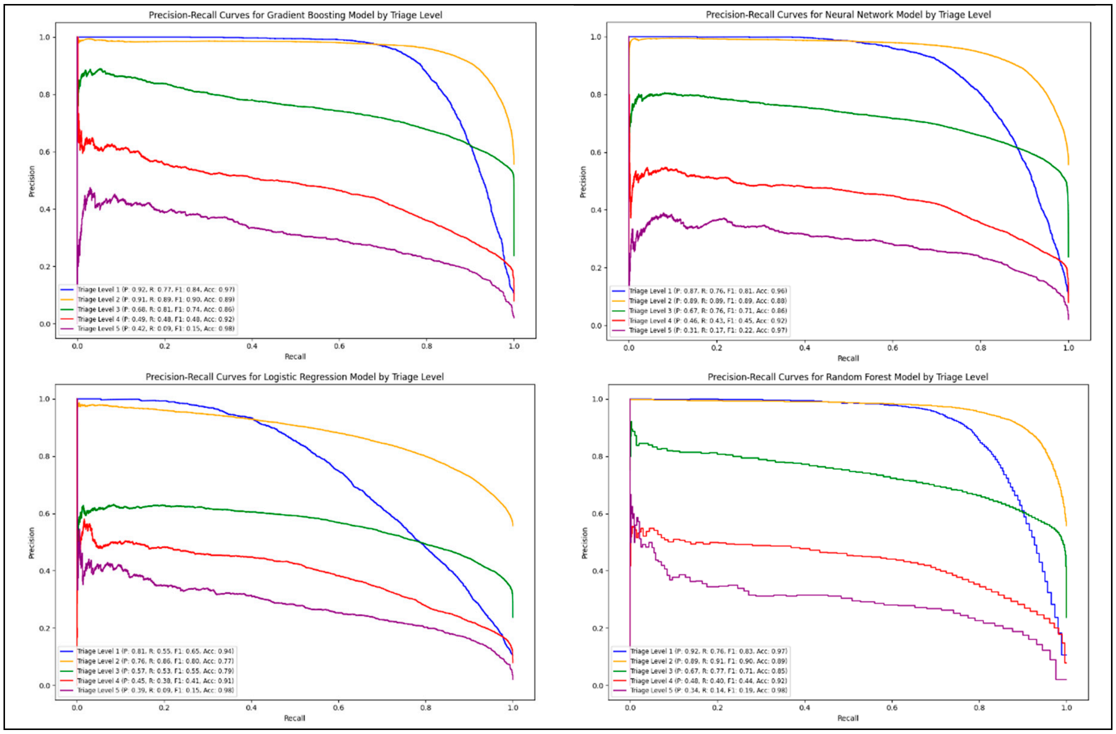
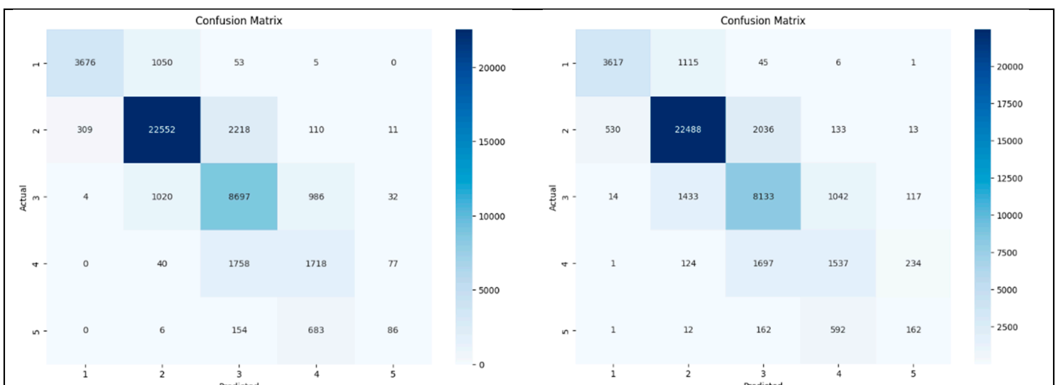


Figure 1. Comparison between models of the precision recall curves by triage level .

Confusion matrices for each model illustrated the distribution of true positive, false positive, true negative, and false negative predictions across the five ESI levels. The gradient boosting and neural network models performed better in correctly classifying higher acuity levels (ESI levels 1 and 2), while logistic regression had more difficulty distinguishing between mid-level acuities (ESI levels 3 and 4) were shown in Figure 2.



Confusion Matrix of Gradient Boosting Confusion Matrix of Neural Network

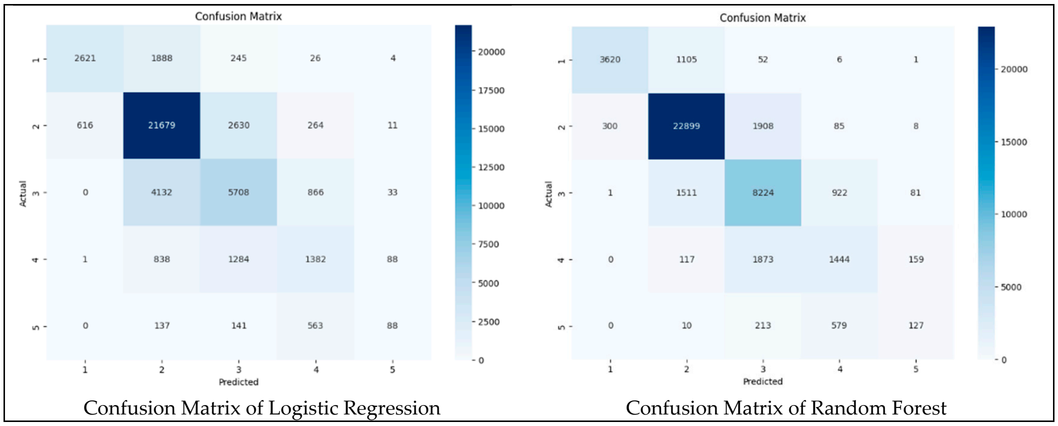


Figure 2. Compare confusion Matrix of different models.

Feature importance analysis identified the most significant predictors for each model. Common important features included pain scale, sex, and mean arterial pressure (MAP). Pain scale emerged as the most critical feature across all models, highlighting its importance in predicting ESI triage levels accurately were shown in Figure 3.

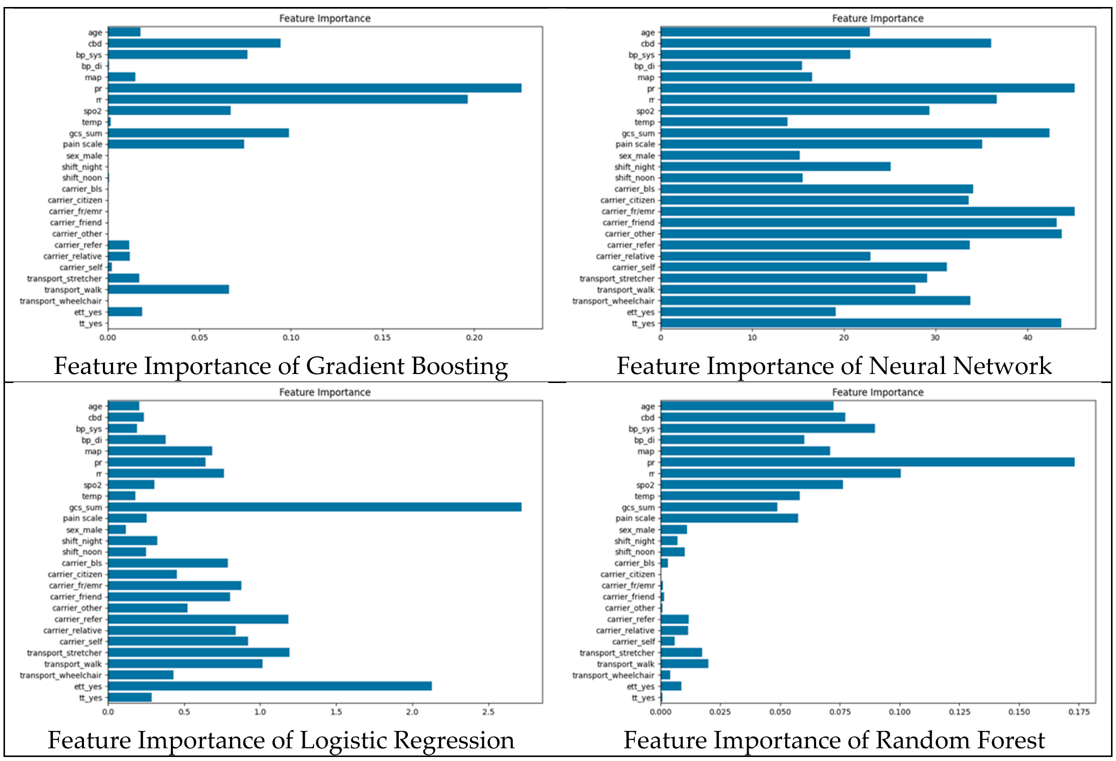


Figure 3. Compare the feature importances of different models.

4. Discussion

Our study compared the performance of logistic regression, gradient boosting, neural network, and random forest models in predicting ESI triage levels for non-traumatic patients in the emergency department of Lampang Hospital. The results demonstrated that gradient boosting and neural network models outperformed logistic regression in terms of accuracy, precision, recall, and F1 score, corroborating findings from previous research (8,9). The superior performance of gradient boosting and neural networks can be attributed to their ability to capture complex nonlinear relationships in the data [7,10].

 doi:10.20944/preprints202408.0428.v1

Precision-recall analysis further highlighted the effectiveness of gradient boosting analysis further highlighted the effectiveness of gradient boosting and neural network models, particularly in correctly identifying true positive cases while minimizing false positives across various ESI levels [11,12]. This is crucial in ED settings where timely and accurate triage decisions can significantly impact patient outcomes and resource allocation.

The confusion matrices revealed that gradient boosting and neural networks were more effective in accurately classifying higher acuity levels (ESI levels 1 and 2) compared to logistic regression [8,13]. This suggests that these models are better suited for identifying patients requiring immediate medical attention, thereby potentially improving ED efficiency and patient care.

Feature importance analysis identified pain scale, sex, and mean arterial pressure (MAP) as key predictors for ESI triage levels across all models [14,15]. These findings align with existing literature emphasizing the value of integrating both objective clinical parameters and subjective patient-reported metrics in predictive modeling [16]. The significance of the pain scale as a predictor highlights its critical role in triage assessments and supports its continued use in ED settings [17].

Despite the promising results, the implementation of ML-based triage systems in EDs faces several challenges. Data privacy concerns, integration with existing ED workflows, and the need for ongoing model validation and updates are significant hurdles that must be addressed [18,19]. Additionally, the variability in patient populations and ED practices necessitates the customization of ML models to ensure their applicability and effectiveness in different settings [20].

Future research should focus on developing strategies to seamlessly incorporate these models into existing triage systems and evaluating their long-term impact on patient outcomes and ED efficiency [21,22]. Moreover, exploring the potential of combining multiple ML models into an ensemble approach could further enhance triage accuracy and reliability. Continuous model training with updated data and real-time validation will be essential to maintain the efficacy of these systems over time [23].

5. Conclusions

This study demonstrates the potential of advanced machine learning models, specifically gradient boosting and neural networks, in improving the accuracy and consistency of Emergency Severity Index (ESI) triage predictions in emergency department settings. By comparing these models with traditional logistic regression, it was evident that gradient boosting and neural networks outperformed logistic regression in terms of accuracy, precision, recall, and F1 score.

The precision-recall analysis highlighted the superior performance of gradient boosting and neural network models in correctly identifying true positive cases while minimizing false positives across various ESI levels. The confusion matrices further confirmed the efficacy of these models, particularly in accurately classifying higher acuity levels (ESI levels 1 and 2), which are crucial for timely medical interventions.

Feature importance analysis underscored the significance of pain scale, sex, and mean arterial pressure (MAP) as key predictors for ESI triage levels. These findings align with existing literature emphasizing the value of integrating both objective clinical parameters and subjective patient-reported metrics in predictive modeling.

The implementation of these advanced machine learning models could significantly enhance the triage process's efficiency and accuracy, ultimately improving patient care and resource management in emergency departments. However, practical integration into clinical workflows will require addressing challenges such as data privacy, continuous model updates, and customization for specific ED settings.

Future research should focus on developing strategies to seamlessly incorporate these models into existing triage systems and evaluating their long-term impact on patient outcomes and emergency department efficiency. This study adds to the growing body of evidence supporting the adoption of machine learning in healthcare, paving the way for more reliable and efficient triage processes in emergency settings.

Supplementary Materials: The following supporting information can be downloaded at: 10.6084/m9.figshare.26232104, Table 1: Baseline characteristics; Table 2: Comparative the overall performance of different models; Figure 1: Comparison between models of the precision recall curves by triage level; Figure 2: Compare confusion Matrix of different models; Figure 3: Compare the feature importances of different models

Author Contributions: Conceptualization, Dr. Noppadon Seesuwan; methodology, Dr. NoppdonSeesuwan; software, Dr. Noppadon Seesuwan, Dr. Auswin Rojanasumapong; validation, Dr. Noppadon Seesuwan, Dr. Auswin Rojanasumapong, Napat Phetkub; formal analysis, Dr. Noppadon Seesuwan, Dr. Napat Phetkub ; investigation, Dr. Noppadon Seesuwan, Dr. Auswin Rojanasumapong; resources, Nipada Chuenjai; data curation, Nipada Chuenjai; writing—original draft preparation, Dr. Noppadon Seesuwan; writing—review and editing, Dr. Napat Phetkub; visualization, Dr. Noppadon Seesuwan; supervision, Napat Phetkub; project administration, Napat Phetkub. All authors have read and agreed to the published version of the manuscript." Please turn to the CRediT taxonomy for the term explanation. Authorship must be limited to those who have contributed substantially to the work reported.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, The Belmont Report, CIOMS Guideline and International Conference on Harmonization in Good Clinical Practice (ICH_GCP) and approved by the Ethics Committee of Lampang Hospital (protocol code EC 093/67 and date of approval 24 July 2024).

Informed Consent Statement: Not applicable.

Data Availability Statement: The original data presented in the study are openly available in FigShare at 10.6084/m9.figshare.26232080

Acknowledgments: The authors would like to thank the staff of the emergency department at Lampang Hospital for their support and assistance in data collection.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Gilboy N, Tanabe T, Travers D, Rosenau AM. Emergency Severity Index (ESI): A Triage Tool for Emergency Department Care, Version 4. Implementation Handbook 2012 Edition. AHRQ; 2011.
2. McHugh M, Tanabe P, McClelland M, Khare RK. More patients are triaged using the Emergency Severity Index than any other triage acuity system in the United States. Acad Emerg Med. 2012;19(1):106-109. doi: 10.1111/j.1553-2712.2011.01240.x. Epub 2011 Dec 23. PMID: 22211429.
3. Hinson J, Martinez D, Cabral S, George K, Whalen M, Hansoti B, Levin S. Triage Performance in Emergency Medicine: A Systematic Review. Ann Emerg Med. 2018;74:10.1016/j.annemergmed.2018.09.022.
4. Kolasa K, Admassu B, Holownia M, Kędzior K, Poirrier JE, Perni S. Systematic Reviews of Machine Learning in Healthcare: A Literature Review. 2023. doi: 10.21203/rs.3.rs-3035548/v1.
5. Delahanty RJ, Alvarez J, Flynn LM, Sherwin RL, Jones SS. Development and Evaluation of a Machine Learning Model for the Early Identification of Patients at Risk for Sepsis. Ann Emerg Med. 2019;73(4):334-344. doi: 10.1016/j.annemergmed.2018.11.036.
6. Habehh H, Gohel S. Machine Learning in Healthcare. Curr Genomics. 2021;22(4):291-300. doi: 10.2174/1389202922666210705124359. PMID: 35273459; PMCID: PMC8822225.
7. Johnson KW, Torres Soto J, Glicksberg BS, Shameer K, Miotto R, Ali M, Ashley E, Dudley JT. Artificial Intelligence in Cardiology. J Am Coll Cardiol. 2018;71(23):2668-2679. doi: 10.1016/j.jacc.2018.03.521. PMID: 29880128.
8. Raita Y, Goto T, Faridi MK, Brown DFM, Camargo CA Jr, 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. PMID: 30795786; PMCID: PMC6387562.
9. Hong WS, Haimovich AD, Taylor RA. Predicting hospital admission at emergency department triage using machine learning. PLoS One. 2018;13(7). doi: 10.1371/journal.pone.0201016. PMID: 30028888; PMCID: PMC6054406.
10. Fei Gao, Baptiste Boukebous, Pozzar Mario, Alaoui Enora, Sano Batourou, Sahar Bayat-Makoei. Predictive Models for Emergency Department Triage using Machine Learning: A Review. Obstetrics and Gynecology Research. 2022;5:107-121.
11. Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. N Engl J Med. 2019;380(14):1347-1358. doi: 10.1056/NEJMra1814259. PMID: 30943338.
12. Beam AL, Kohane IS. Big Data and Machine Learning in Health Care. JAMA. 2018;319(13):1317-1318. doi: 10.1001/jama.2017.18391. PMID: 29532063.

13. National Institute of Emergency Medicine. Emergency patient care triage and prioritization criteria set by NIEMS 2013 [Internet]. 2013 [cited 2024 April 4]. Available from:https://www.niems.go.th/1/Ebook/Detail/272?group=21
14. Breiman L. Random Forests. Mach Learn. 2001;45:5-32. doi: 10.1023/A:1010933404324.
15. LeCun Y, Bengio Y, Hinton G. Deep Learning. Nature. 2015;521:436-444. doi: 10.1038/nature14539.
16. Freund Y, Schapire RE. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. J Comput Syst Sci. 1997;55(1):119-139. doi: 10.1006/jcss.1997.1504.
17. Kumar R, Indrayan A. Receiver operating characteristic (ROC) curve for medical researchers. Indian Pediatr. 2011;48(4):277-287. doi: 10.1007/s13312-011-0055-4. PMID: 21532099.
18. Chen TQ, Guestrin C. Xgboost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016 Aug 13-17; San Francisco. p. 785-794. doi: 10.1145/2939672.2939785.
19. Esteva A, Kuprel B, Novoa R, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542:115-118. doi: 10.1038/nature21056.
20. Davis J, Goadrich M. The relationship between Precision-Recall and ROC curves. In: Proceedings of the 23rd International Conference on Machine Learning (ICML '06). New York, NY: Association for Computing Machinery; 2006. p. 233-240. doi: 10.1145/1143844.1143874.
21. Saito T, Rehmsmeier M. The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets. PLoS One. 2015;10(3). doi: 10.1371/journal.pone.0118432.
22. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. Advances in neural information processing systems. 2017;30:4765-4774.
23. Duda R, Hart P, Stork D. Pattern Classification. 2nd ed. New York: Wiley; 2001.