The Reliability of an Artificial Intelligence Tool, ‘Decision Trees’, in Emergency Medicine Triage

Alp Giray AYDIN (✉ dralp07@gmail.com)
Koc University Medical Faculty  https://orcid.org/0000-0002-3093-6120

Oktay Eray
Akdeniz Universitesi School of Medicine Department of Emergency Medicine

Ali Vefa SAYRAC
Chief Physician Atatürk Public Hospital

Alten OSKAY
Pamukkale University School of Medicine Department of Emergency Medicine

Umit Deniz ULUSAR
Akdeniz University Engineering Faculty Department of Computer Engineering

Original Research

Keywords: emergency department, triage, artificial intelligence, decision trees

DOI: https://doi.org/10.21203/rs.3.rs-127447/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

Objective: Overcrowding is a challenge for emergency departments throughout the world. Triage systems categorize the patients based on medical emergencies in order to avoid the malpractices. The present study aimed to test the validity of an artificial intelligence tool, ‘Decision Trees’, in emergency medicine triage.

Methods: This prospective, cross-sectional, clinical study was conducted in an emergency department of a tertiary care hospital. A total of 1999 patients over 18 years were included into the study. The triage staff were trained before the study with the Australasian Triage Scale. Two independent observers rate the ultimate triage category of study patients. A new algorithm by ‘Decision Trees’ was constructed at the end of the study.

Results: The mean age of the study patients were 41.1±17.2 and 49.1 % of them (n=981) were male. There were 867 patients (43.3%) with triage category of five and 14 (0.7%) patients with triage category one. The most common clinical descriptors of the patients were minimal pain with no high risk features 20.5% of them (n=409) and minor symptoms of low risk conditions 18.1% of them (n=362). There was an excellent consistency between two independent observers (kappa value: 0.997. The new algorithm by ‘Decision Trees’ rated wrong in only one patient. The accuracy rate was 99.9%. The consistency between ATS and ‘Decision Trees was excellent (kappa value: 0.999). There was average consistency between physicians and paramedics. (kappa value: 0.541).

Conclusion: Decision trees as an artificial intelligence model should be used for producing practical triage algorithms as a decision support tool in emergency departments.

Introduction

Overcrowding is one of the most important issues of emergency departments (ED) in recent years. Increase in overcrowding leads to delays in the evaluation and the treatments of patients that results with a negative impact on patient satisfaction and the quality of the medical care (1, 2). The requirement for triage was arisen from to the necessity of selecting patients need urgent care.

Triage systems have had a common use in so many countries in recent years. The triage scales performed by the medical stuff categorizes the patients by using various medical parameters and presentation symptoms (3).

Artificial intelligence, which models the logic of human brain, has been used since 1969 in some fields of medicine with promising results (4). There are so many techniques for artificial intelligence, which have the capacity of solving problems, and the usage of these systems in medicine makes sense.

This study aimed to reveal the validity of an artificial intelligence method, ‘Decision Trees’, in triaging the patients presented to ED.
Material And Method

This prospective cross sectional trial was conducted in an ED of a tertiary care hospital with an annually census of 73000 patients between 1–10 July in 2011. The local ethical committee approved the study.

Patients presented to the ED with any complaints over 17 years old included into the study. Patients under 18 years old and denied to give inform consent was excluded from the study.

Eight paramedics and one nurse working in the triage area were trained for Australia Triage Scale (ATS) (Supplement) before the study. During the training, the categories of ATS were defined firstly and clinical scenarios composed the second stage. In the third stage, a practical application was implemented with 100 patients in the triage area.

Blood pressure, body temperature, oxygen saturation, heart and respiration rate, age, gender, presenting symptom, co-morbidities, Glasgow Coma Scale (GCS) score, verbal pain score that classifying the patients as serious pain, average pain, less pain and no pain and the triage category according to ATS were recorded to the study form. Blood pressure, pulse oxymeter and pulse rate were measured by a monitor. Body temperature was measured by temporal route (Exergen Temporal Artery Thermometry).

Triage Categories for ATS

Category 1: Patients who need resuscitation and admitted to the resuscitation area immediately

Category 2: Patients with a life-threatening condition or risky for any limb. These patients were admitted to the telemetry unit or resuscitation area within 10 minutes.

Category 3: Patients with pathology of potential to progress life-threatening conditions or urgent interventions. These patients were admitted to the ED in 30 minutes.

Category 4: Patients who have pathologies that may take care in one and two hours. These patients admitted to ED within 60 minutes.

Category 5: Patients with non-urgent complaints and admitted to the ED in 120 minutes.

The details of ATS have been displayed in Supplement.

A chart defining ATS was posted on the triage area for triage stuff. Patients waiting for ED admission re-triaged with 30 minutes intervals for any change in triage category.

Two independent senior residents at the end of the study determined the ultimate triage category. An associated professor on emergency medicine evaluated the patients if there was any inconsistency between two observers.
Statistical Analysis

‘Decision trees’ as the artificial intelligence model was used in the present study. ‘Mathlab in classregtree method’ was used for decision tree analysis. This method generates decision trees from the upper root point to down. There is a separation parameter and a condition in every branching point. Separation criteria providing most information were used in the algorithm (Gini’s diversity index). A branch was formed in the right and left side of every criteria and value. The values below the selected variable and value are usually localized at the left sight. The algorithm continues to construct the peripheral branches until there is little point. The smallest point number in branches for calculation was used, 1. The other analysis of the study was performed by MedCalc. Numerical data was presented as mean ± standard deviation or median (interquartile range/min-max) and categorical data as rates. The consistency between decision trees, paramedic and physicians were displayed by kappa value.

Results

Two thousand patients who gave inform consent were included into the study. One patient with lack of data in the study form excluded from the study and 1999 patients included into the nal analysis. The mean age of the study subjects was 41.1 ± 17.2 and 49% (n = 981) of them were male. The mean systolic blood pressure was 126 ± 23 mmHg and diastolic pressure was 72 ± 13. The median respiratory rate was 18/min (min-max:0–48), median pulse rate was 85/min (min-max: 0-200) and mean oxygen saturation 96 ± 8. The most frequent co-morbidities in study patients were hypertension (17.4% (n = 347)) and diabetes mellitus (10.8% (n = 215)).

There were 867 patients (43.3%) with triage category of five and 14 (0.7%) patients with triage category one (Fig. 1). There were 49 patients (2.5%) with severe pain, 342 patients (17.1%) with average pain, 974 patients (48.7%) with less pain and 634 patients (31.7%) with no pain.

The median GCS score was 15 (IQR: 15–15, min-max: 3–15). The demographics of study patients were presented in Table 1.
Table 1
Demographics of study patients

| Variable                                | Value                  |
|-----------------------------------------|------------------------|
| Age (mean ± sd)                         | 41.1 ± 17.2            |
| Gender                                  | 981/1018 49.1/50.9     |
| Systolic Blood Pressure (mean ± sd)     | 126.4 ± 22.8           |
| Diastolic Blood Pressure (mean ± sd)    | 72.2 ± 12.9            |
| Body temperature (mean ± sd)            | 36.3 ± 1.5             |
| Pulse rate (median (IQR))               | 85 min-max: 0-200      |
| Oxygen saturation (mean ± sd)           | 96.5 ± 7.9             |
| Respiration rate (median (IQR))         | 18 min-max: 0–48       |
| Diabetes Mellitus                       | 215 (10.8)             |
| Hypertension                            | 347 (17.4)             |
| Cardiac Failure                         | 89 (4.5)               |
| Coronary Artery Disease                 | 122 (6.1)              |
| Malignancy                              | 93 (4.7)               |
| Others                                  | 87 (4.4)               |
| GCS score (median (IQR))                | 15 (min-max: 3–15)     |
| Verbal Pain Score                       |                        |
| Severe                                  | 49 (2.5)               |
| Average                                 | 342 (17.1)             |
| Less pain                               | 974 (48.7)             |
| No pain                                 | 578 (28.9)             |
| Not evaluated                           | 56 (2.8)               |

Abbreviations: sd: standard deviation; IQR: interquartile range; GCS: Glasgow Coma Scale

The most frequent clinical criteria according to ATS was minimal pain without high risk (409 (20.5%)), minor symptoms with stable condition (362(18.1%)), average pain with a risky condition (168 (8.4%)), diarrhea and vomiting without dehydration (158 (7.9%)). Eight patients (0.4%) with cardiac arrest, one patient (0.1%) with respiratory arrest and four patients (0.2%) with a GCS score of under 9 were admitted
to the ED during the study period. The frequency of clinical criteria according to ATS was displayed in Fig. 2 and Table 2a and 2b.
Table 2a
Clinical descriptors determined according to ATS. The identifier numbers in parentheses next to clinical descriptors refers to the specified numbers in ‘Decision Trees’ algorithm.

| Chief Complaint                                          | Number of Patients | Percent |
|----------------------------------------------------------|--------------------|---------|
| Cardiac arrest (1)                                       | 8                  | 0.4     |
| Respiratory arrest (2)                                   | 1                  | 0.1     |
| Respiratory rate < 10/min (4)                            | 1                  | 0.1     |
| Unresponsive or responds to pain only (GCS < 9) (7)      | 4                  | 0.2     |
| Extreme respiratory distress (12)                        | 16                 | 0.8     |
| Circulatory compromise (13)                              | 21                 | 1.1     |
| - clammy or mottled skin, poor perfusion                 |                    |         |
| - heart rate < 50 or > 150                               |                    |         |
| - hypotension with haemodynamic effects                   |                    |         |
| - severe blood loss                                       |                    |         |
| Chest pain of likely cardiac nature (14)                 | 35                 | 1.8     |
| Very severe pain- any cause (15)                          | 3                  | 0.2     |
| Drowsy, decreased responsiveness any cause (GCS < 13) (16)| 9                  | 0.5     |
| Acute hemiparesis/dysphasia (17)                         | 24                 | 1.2     |
| Acid or alkali splash to eye- requiring irrigation (20)   | 1                  | 0.1     |
| Major multi trauma (21)                                  | 8                  | 0.4     |
| Severe localized trauma- major fracture, amputation (22)  | 3                  | 0.2     |
| High-risk history (23)                                   | 16                 | 0.8     |
| - significant sedative or other toxic ingestion           |                    |         |
| - severe pain suggesting PE, AAA, ectopic pregnancy       |                    |         |
| Behavioural/Psychiatric (24)                             | 3                  | 0.2     |
| - violent or aggressive                                   |                    |         |
| - immediate threat to self or others                      |                    |         |
| - severe agitation or aggression                          |                    |         |
| Severe hypertension (25)                                 | 2                  | 0.1     |

Abbreviations: GCS: Glasgow Coma Scale; LOC: Loss of consciousness
| Chief Complaint                                                                 | Number of Patients | Percent |
|--------------------------------------------------------------------------------|--------------------|---------|
| Moderately severe blood loss- any cause (26)                                   | 7                  | 0.4     |
| Moderate shortness of breath (27)                                              | 33                 | 1.2     |
| SaO2 90–95% (28)                                                                | 9                  | 0.5     |
| Seizure (now alert) (29)                                                        | 7                  | 0.4     |
| Any fever if immunosupressed eg oncology patient, steroid Rx (30)               | 24                 | 1.2     |
| Persistent vomiting (31)                                                        | 20                 | 1       |
| Dehydration (32)                                                                | 3                  | 0.2     |
| Head injury with short LOC (now alert) (33)                                     | 1                  | 0.1     |
| Moderately severe pain-any cause-requiring analgesia (34)                       | 25                 | 1.3     |
| Chest pain likely non-cardiac and abdominal pain without high risk features (severe or patient age > 65 ) (35) | 35                 | 1.8     |
| Moderate limb injury- deformity, severe laceration, crush (36)                  | 9                  | 0.5     |

*Abbreviations: GCS: Glasgow Coma Scale; LOC: Loss of concious*
| Chief Complaint                                                                 | Number of Patients | Percent |
|--------------------------------------------------------------------------------|--------------------|---------|
| Limb- altered sensation, acutely absent pulse and trauma- high-risk history    | 7                  | 0,4     |
| with no other risk features (37)                                               |                    |         |
| Behavioural/Psychiatric (40)                                                   | 13                 | 0,7     |
| - very distressed, risk of self-harm                                           |                    |         |
| - acutely psychotic or thought disordered                                      |                    |         |
| - situational crisis, deliberate self harm                                     |                    |         |
| - agitated/ withdrawn                                                          |                    |         |
| - potentially agressive                                                        |                    |         |
| Mild haemorrhage (41)                                                          | 81                 | 4,1     |
| Chest injury without rib pain or respiratory distress (43)                    | 7                  | 0,4     |
| Difficulty swallowing, no respiratory distress (44)                           | 2                  | 0,1     |
| Moderate pain, some risk features (46)                                         | 168                | 8,4     |
| Vomiting or diarrhoea without dehydration (47)                                 | 158                | 7,9     |
| Eye inflammation or foreign body- normal vision (48)                          | 75                 | 3,8     |
| Minor limb trauma- sprained ankle, possible fracture, uncomplicated laceration| 162                | 8,1     |
| requiring investigation or intervention- Normal vital signs, low/moderate pain |                    |         |
| (49)                                                                           |                    |         |
| Nonspecific abdominal pain (50)                                                | 121                | 6,1     |
| Behavioural/Psychiatric (51)                                                   | 10                 | 0,5     |
| - Semi-urgent mental health problem                                            |                    |         |
| - Under observation and/or no immediate risk to self or others                 |                    |         |
| Minimal pain with no high risk features (52)                                   | 409                | 20,5    |
| Low risk history and now asymptomatic (53)                                     | 9                  | 0,5     |
| Stable minor symptoms (54)                                                     | 362                | 18,1    |
| Minor symptoms of existing stable illness (55)                                 | 1                  | 0,1     |
| Minor symptoms of low-risk conditions (56)                                     | 11                 | 0,6     |
There was an excellent consistency between two independent observers (kappa value: 0.997). Only five patients were rated different between two observers (0.2%).

A new triage algorithm was constructed by decision trees (Fig. 3). The new algorithm by ‘Decision Trees’ rated wrong in only one patient. The accuracy rate was 99.9%. The consistency between ATS and ‘Decision Trees’ was excellent (kappa value: 0.999).

The impact of parameters in determining triage categorization was displayed in Fig. 4.

There was average consistency between physicians and paramedics. (kappa value: 0.541) (Fig. 5).

Discussion

The term artificial intelligence was firstly defined by John McCarthy as a science and engineering of producing clever machines, particularly clever computer software (5). Artificial intelligence is the performing of behaviors, called intelligence when conducted by humanity, by machines. It is composed of methods that model the imagination of human being and the work form of human brain.

The ability of artificial intelligence models in revealing the relations in a data set is used for forecasting the diagnosis, treatments and prognosis in many clinical scenarios (6). There are many techniques of artificial intelligence in order to be used for various clinical problems. The potency of these techniques in investigating the diseases and treatment modalities is promising.

‘Decision Trees’ is one of the artificial intelligence models that used as a decision support system. Decision Trees constructs a prediction model between a patient and disease with the parameters that we determine. Information is presented as a tree composed of steps; decision step, probability step and conclusion step.

The fundamentals of artificial intelligence are based on 1965 and defining applicability of fuzzy clusters in medicine in 1969 leads so researches about it. Successful researches in cardiology in the first years were followed by new investigations in radiology and other areas of medicine (7–12). Artificial intelligence models are used in various areas of medicine nowadays (13–16).
Besides the usage of artificial intelligence in many areas of medicine, models of artificial intelligence have been constructed also for emergency medicine. Artificial neural network, a kind of artificial intelligence, was tested to predict the acute coronary syndromes in emergency department, which found to have a reasonable performance with AUC value of 0.97, 0.93 and 0.95 from different hospitals. (17). Bektas et al. used artificial neural network to predict the craniocervical junction injuries in trauma patients and reported the AUC value as 0.912 (18).

In a recent trial by Azeez et al., validity of artificial intelligence in triage system was tested by producing prediction models with retrospective chart data by using artificial neural network and neuro-fuzzy system (19). The authors reported the accuracy of ANN and neuro-fuzzy system in test-set as 96.7 and 94, respectively. Although the specificities of two methods in prediction the triage category accurately over 90%, the sensitivities are lower particularly in neuro-fuzzy models except in one of the prediction models in both methods.

The present study showed that ‘Decision Trees’ may be used to construct highly predictive and practical algorithms for implanting triage. The model with decision trees in the present trial rated wrong only one among 1999 patients. These methods are so flexible that you can change and produce new algorithms with different variables that the investigator chooses.

Actually, the results of this trial do not claim that artificial intelligence should substitute the paramedics or triage nurses. However, they should be used as decision aid tools for triaging patients, particularly for beginner triagers and also in crowded EDs which needs rapid decision.

**Conclusion**

Decision trees as an artificial intelligence model should be used for producing practical triage algorithms as a decision support tool in emergency departments.

**Declaration**

**Financial support and conflicts of interests**

None declared

**References**

1. Shiber JR, Longley MB, Brewer KL. Hyper-use of the ED. Am J Emerg Med 2009; 27: 588-94.
2. Pines JM, Prabhu A, McCusker CM, Hollander JE. The effect of ED crowding on education. Am J Emerg Med 2010; 28 (2): 217-20.
3. Dikme O, Atilla R, Colak Oray N, Saracoglu G, Girgin M, Akyol P. Dokuz Eylul Triage System: A five category triage algorithm, reliability and validity study. Anatolian J Emerg Med 2019; 2(3): 1-7.
4. Serhatlioğlu S, Hardalac F. Yapay Zeka Teknikleri ve Radyolojiye Uygulanması. Fırat Tıp Dergisi 2009; 14 (1): 1-6.

5. McCarthy J. What is artificial intelligence? Computer Science Department, Stanford University. Available from: http://www.formal.stanford.edu/ jmc/whatisai.pdf

6. Demirhan A, Kılıç YA, Güler I. Tıpta Yapay Zeka Uygulamaları. Yoğun Bakım Dergisi 2010; 9(1): 31-41.

7. Zadeh LA. Biological application of the theory of fuzzy sets and systems on Biocybernetics of the Central Nervous System, Proc Int Sym 1969; 199-212.

8. Kalmanson D, Stegall HF. Cardiovascular investigations and fuzzy set theory. American Journal of Cardiology 1975; 35: 80-4.

9. Guo Z, Durand LG, Allard L, Cloutier G, Lee HC, Langlois YE. Cardiac Doppler blood flow signal analysis. Part II:The timefrequency distribution by using autoregressive modeling. Med Biol Eng Comput 1993; 31: 242-8.

10. Jain R, Mazumdar J, Moran W. Application of fuzzy classifier system to coronary artery disease and breast cancer. Australasian Physical Engineering Sciences in Medicine 1998; 21(3): 141-7.

11. Guler D, Hardalac F, Bansci, N. Application of FFT analyzed Cardiac Doppler Signals To Fuzzy Algorithm. Computers in Biology and Medicine 2002; 32: 435-44.

12. Guler D, Hardalac F, Ergu, U, Barıscı N. Classification of Aorta Doppler signals using variable coded-hierarchical genetic fuzzy system. Expert Systems with Applications 2004; 26: 321-33.

13. Letourneau S, Jensen L. Impact of a decision tree on chronic wound care. J.Wound Ostomy Continence Nurs 1998; 25: 240-47.

14. Gülkesen KH, Köksal IT, Özdem S, Saka O. Prediction of prostate cancer using decision tree algorithm. Turk J Med Sci 2010; 40(5): 681-6.

15. Eken C, Bilge U, Kartal M, Eray O. Artificial neural network, genetic algorithm, and logistic regression applications for predicting renal colic in emergency settings. Int J Emerg Med 2009; 2 (2): 99-105

16. Liu Y, Wei LQ, Li GQ, Lv FY, Wang H, Zhang YH, et A decision-tree model for predicting extubation outcome in elderly patients after a successful spontaneous breathing trial. Anesth Analg 2010; 111(5): 1211-8.

17. Harrison RF, Kennedy RL. Artificial neural network models for prediction of acute coronary syndromes using clinical data from the time of presentation. Ann Emerg Med. 2005 Nov;46(5):431-9.

18. Bektaş F, Eken C, Soyuncu S, Kilicaslan I, Cete Y. Artificial neural network in predicting craniocervical junction injury: an alternative approach to trauma patients. Eur J Emerg Med. 2008 Dec;15(6):318-23. doi: 10.1097/MEJ.0b013e3282fce7af.

19. Azeez D, Mohd Ali MA, Gan KB, et al. Comparison of adaptive neuro-fuzzy inference system and artificial neural networks model to categorize patients in the emergency department. Springer plus. 2013;2:416.
Figures

Figure 1
Ultimate triage scores of study patients.

Figure 2
The frequencies of clinical criteria in ATS.
Figure 3

The algorithm of ‘Decision Trees’.

Figure 4

The impact of parameters in decision of triage categorization.
Figure 5

Consistency between physicians and paramedics.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplement.docx