Artificial Intelligence In Medicine 102 (2020) 101762

Contents lists available at ScienceDirect

Artificial Intelligence In Medicine

journal homepage: www.elsevier.com/locate/artmed

Clinical Decision Support Systems for Triage in the Emergency Department using Intelligent Systems: a Review

Marta Fernandes[a,*], Susana M. Vieira[a], Francisca Leite[b], Carlos Palos[c], Stan Finkelstein[d], João M.C. Sousa[a]

[a] IDMEC, Instituto Superior Técnico, Universidade de Lisboa, Portugal
[b] Hospital da Luz Learning Health, Portugal
[c] Hospital Beatriz Ângelo, Luz Saúde, Portugal
[d] Institute for Data, Systems and Society, Massachusetts Institute of Technology, Massachusetts

ARTICLE INFO

Keywords:
Triage
CDSS
EHR
Machine learning
Critical care

ABSTRACT

Motivation: Emergency Departments' (ED) modern triage systems implemented worldwide are solely based upon medical knowledge and experience. This is a limitation of these systems, since there might be hidden patterns that can be explored in big volumes of clinical historical data. Intelligent techniques can be applied to these data to develop clinical decision support systems (CDSS) thereby providing the health professionals with objective criteria. Therefore, it is of foremost importance to identify what has been hampering the application of such systems for ED triage.
Objectives: The objective of this paper is to assess how intelligent CDSS for triage have been contributing to the improvement of quality of care in the ED as well as to identify the challenges they have been facing regarding implementation.
Methods: We applied a standard scoping review method with the manual search of 6 digital libraries, namely: ScienceDirect, IEEE Xplore, Google Scholar, Springer, MedlinePlus and Web of Knowledge. Search queries were created and customized for each digital library in order to acquire the information. The core search consisted of searching in the papers' title, abstract and key words for the topics "triage", "emergency department"/"emergency room" and concepts within the field of intelligent systems.
Results: From the review search, we found that logistic regression was the most frequently used technique for model design and the area under the receiver operating curve (AUC) the most frequently used performance measure. Beside triage priority, the most frequently used variables for modelling were patients' age, gender, vital signs and chief complaints. The main contributions of the selected papers consisted in the improvement of a patient's prioritization, prediction of need for critical care, hospital or Intensive Care Unit (ICU) admission, ED Length of Stay (LOS) and mortality from information available at the triage.
Conclusions: In the papers where CDSS were validated in the ED, the authors found that there was an improvement in the health professionals' decision-making thereby leading to better clinical management and patients' outcomes. However, we found that more than half of the studies lacked this implementation phase. We concluded that for these studies, it is necessary to validate the CDSS and to define key performance measures in order to demonstrate the extent to which incorporation of CDSS at triage can actually improve care.

1. Introduction

The growing demand for emergency services, combined with the priority sorting due to patient's acuity, results in long waiting times for patients. Waiting times have a significant impact on patient mortality, morbidity with readmission in less than 30 days, number of pre-Intensive Care Units (ICU) resuscitation, length of stay (LOS), patient satisfaction and costs [1–7]. The outcome of patients' medical treatment is time-sensitive, therefore the sooner the treatment is rendered, the better the outcome [3–7].

The first point where the patient acuity state is evaluated takes place at the triage stage in the emergency department (ED). Triage systems

* Corresponding author.
E-mail addresses: marta.fernandes@tecnico.ulisboa.pt (M. Fernandes), susana.vieira@tecnico.ulisboa.pt (S.M. Vieira), francisca.leite@hospitaldaluz.pt (F. Leite), Carlos.Palos@hbeatrizangelo.pt (C. Palos), snfinkel@mit.edu (S. Finkelstein), jmsousa@tecnico.ulisboa.pt (J.M.C. Sousa).

https://doi.org/10.1016/j.artmed.2019.101762
Received 22 February 2019; Received in revised form 29 August 2019; Accepted 7 November 2019
0933-3657/

are used by health professionals to assign priority levels to patients, based on their urgency of treatment. However, there may be critical patients presenting symptoms not easily recognized as indicators of criticality. If not identified in due time, these patients have to wait for a long time for medical observation, which results in an increased risk of morbidity. Therefore, the need for an efficient triage system to assist the health professional in taking a timely and correct decision becomes of vital importance.

Decision-making support systems (DMSS), or simply decision support systems (DSS), are information systems designed to assist decision-makers and interactively support all phases of a user's decision-making process [8]. With the accentuated progress of Artificial Intelligence (AI) in the 1980s, AI tools were incorporated in DSS to increase the impact of management support. This led to the emergence of intelligent decision support systems (i-DSS) as a sub-discipline of DSS research [9]. A particular technology used within i-DSS research is machine learning, which allows DSS to obtain new knowledge or to adapt to the user or changing environment [9]. An i-DSS extends traditional DSS by incorporating techniques to supply intelligent behaviours and utilizing the power of modern computers to support and enhance decision making [10–12]. The i-DSS may, for example, "respond quickly and successfully to new data and information without human intervention, deal with perplexing and complex situations, learn from previous experience, apply knowledge to understand the environment, recognize the relative importance of different elements in the decision, incorporate the knowledge of domain experts, recommend action, and/or act on behalf of the human (by a predefined authorization of the decision-maker)" [8].

In the clinical setting, DSS are denominated clinical decision support systems (CDSS) and provide clinicians, staff and patients with knowledge, patient-specific information and recommendations. CDSS are usually used for addressing clinical needs, such as ensuring accurate diagnoses, screening in a timely manner for preventable diseases, averting adverse drug events [13] or pain management [14]. However, CDSS can also potentially improve efficiency, reduce costs and patient inconvenience. The aim of such systems is not to replace the decision-makers – clinicians, patients and health organizations – but to provide relevant knowledge and support in their decision-making [15,16].

With the promotion of the implementation of electronic health records (EHRs) [17], there has been a slow but increasing adoption of health information technology worldwide. The adaptation and extra time required in the initial learning stages by health professionals as well as implementation costs were barriers to the progression of the full implementation of the EHR systems. This has led to the adoption of CDSS technology to improve costs and quality of care [18–21]. CDSS designed recurring to intelligent techniques are currently being used for several medical applications [14]. Interpreting clinical data to classify patients in a timely manner is vital in the ED setting, with impacts on costs, efficiency and quality of care. Therefore, there is potential for improvement of ED operations using AI [22]. The implementation of intelligent CDSS for ED triage however presents challenges which will be addressed in this paper.

In this paper, we aim to present a scoping review of CDSS designed with intelligent techniques for ED triage. We assess how these systems have been assisting clinical decisions that resulted in improved quality of care at the ED triage. We consider that a CDSS contributed to ED triage if the results achieved with this system overcame or complemented the ones provided by the triage system implemented at the time in the ED. We also assess which variables, intelligent techniques and performance measures were used in the design and evaluation of the triage CDSS.

Section 2 describes the methodology adopted to achieve the paper's goals. The research methodology consists of a scoping review method, with a manual search of specific digital libraries, which contain papers within the field of intelligent systems, DSS and ED triage. In Section 3, the search results are presented and the main contributions, limitations and future work are addressed and discussed in Section 4. Section 5 offers the conclusions of this review.

2. Methods

The research technique used to perform the scoping review was based on the guidelines proposed by Kitchenham [23,24]. This technique consisted of a literature review of CDSS which were designed based on intelligent techniques and that contributed to assisting health professionals in their decision-making at the ED triage stage.

2.1. Search strategy

The strategy for this review was composed by the following steps:

1) Selection of the data sources for extraction of information.
2) Creation of queries to perform the search in the databases.
3) Collection and summary of the entire gathered information.
4) Outline of the inclusion and exclusion criteria.
5) Data analysis.

2.1.1. Search method

The search method consisted of a manual search performed on 6 digital libraries, namely: ScienceDirect, IEEE Xplore, Google Scholar, Springer, MedlinePlus (PubMed) and ISI Web of Knowledge, as performed in previous studies [14,25–28].

These specific digital libraries were selected because they were known to include studies related to intelligent systems in the healthcare field. It was hypothesized that only after the year of 2004 would we be likely to find published papers combining the two referred topics. It was only from this year on that there were governmental incentives to create EHRs in hospitals [17]. However, the year of publication for the search period was not constrained.

2.1.2. Search terms

Different search queries (SQ) were created for each digital library given that each one has specific features for advanced search, such as the maximum number of input key words. Therefore, the queries had to be customized for each digital library in order to get the required information.

The first generic SQ that was created to search each digital library consisted of searching for the topics "triage", "Emergency Department", "Emergency Room" and concepts within the field of intelligent systems, namely: "machine learning", "modeling", "model", "classification" and "predictive". For ScienceDirect, IEEE Xplore and Pubmed, the terms in the first generic query were searched in the title, abstract and key words of the papers. For Web of Knowledge, these terms were searched as within the"Topic" of the papers, since there is no field tag in this library for specific search of the abstract. Nonetheless, using the"Topic" field tag covers all these fields. In the case of Springer library, we decided that the title should contain the terms "triage" and "Emergency Department", since the advanced search presents limitations regarding field tags. For this library, we identified relevant disciplines, namely "Engineering", "Computer Science", "Statistics" and "Health Informatics" which potentially would yield the search results meeting the search criteria (SQ11 to SQ14 in Table 2). For the cases of ScienceDirect (SQ2 to SQ6 in Table 2) and Web of Knowledge (SQ18 and SQ19 in Table 2), the search was refined by specific Journals related to healthcare or intelligent systems.

In the cases where the number of search results was higher than 1,000, it was necessary to specify exclusion criteria terms that should not appear in the title of the papers. The list of terms is presented in SQ16, SQ18 and SQ19 in Table 2, and includes: "pediatric", "pregnancy", "mental", "psychiatry", "trauma", "sepsis", "chest pain", "epidemy", among others.

We considered Google Scholar to be an addition to the search of the other libraries [26], and we limited the number of results by imposing that the title should contain the terms “triage” and “Emergency Department”/“Emergency Room”. The number of search characters in this digital library is limited, therefore only a few words could be selected to be excluded from the title. From the total number of terms already mentioned as exclusion criteria, we assessed for each term, the ones leading to inferior numbers of search results. In these cases, the title should not contain the terms “trauma”, “sepsis”, “children”, “pediatric” and “chest pain”. When applying the query SQ9 in Table 2, the method for selecting the papers consisted first of reading the title – papers for which the title contained the terms in exclusion criteria where not selected –, and then reading the abstract to assess if the papers satisfied the search criteria.

Although the queries were different, due to the limitation of the advanced search in each digital library, the core search was consistent through all, where the main search terms were “Triage”, “Emergency Department”/“Emergency Room” and those related with intelligent systems.

2.2. Inclusion and exclusion criteria

The inclusion and exclusion (IE) criteria were defined for selecting the relevant studies and filtering the irrelevant ones, which were excluded in the search. The IE criteria applied is presented in Fig. 1.

After applying the queries in the digital libraries, a folder for each one was created on Mendeley. Papers that met the inclusion criteria were saved in a separate folder for the respective library. At the end of this process, all selected papers were joined in a separate folder.

Table 1
Presentation of collected data in the results section.

Data	Presentation
Flow diagram with SQ results from the digital libraries	Fig. 2
Objectives of the study	Fig. 3
Variables used for the development of the triage models	Fig. 5
Intelligent techniques used to develop the triage models	Figures 6 and 7
Performance measures used to evaluate the triage models	Figs. 8 and 9
Key-words	Fig. 4
Main contributions of the selected papers	Subsection 4.1
Limitations and future work of the selected papers	Subsection 4.2
Triage system implemented in the ED case-study	Table 3
Source (journal article or conference paper), reference and results achieved by outcome	Table 4

2.3. Data analysis

The analysis of the data extracted from each study is presented as indicated in Table 1.

3. Results

From the literature search, 62 papers were selected when applying the IE criteria, as presented in Fig. 2. Information regarding the authors, title, publication date and source of these papers is presented in Table 4. As we hypothesized, the papers date back to 2005, following the international governmental incentives for the existence of EHRs.

3.1. Main findings of the review search

The aim of the majority of the selected papers was to develop

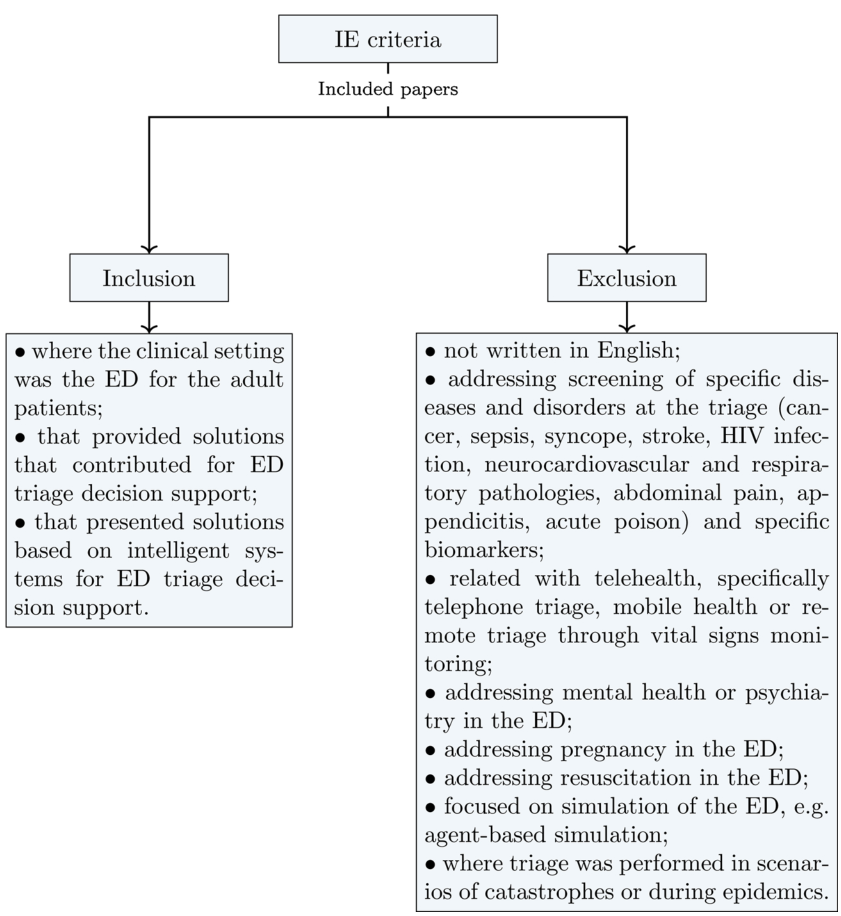

Fig. 1. Inclusion and exclusion (IE) criteria for papers selection.

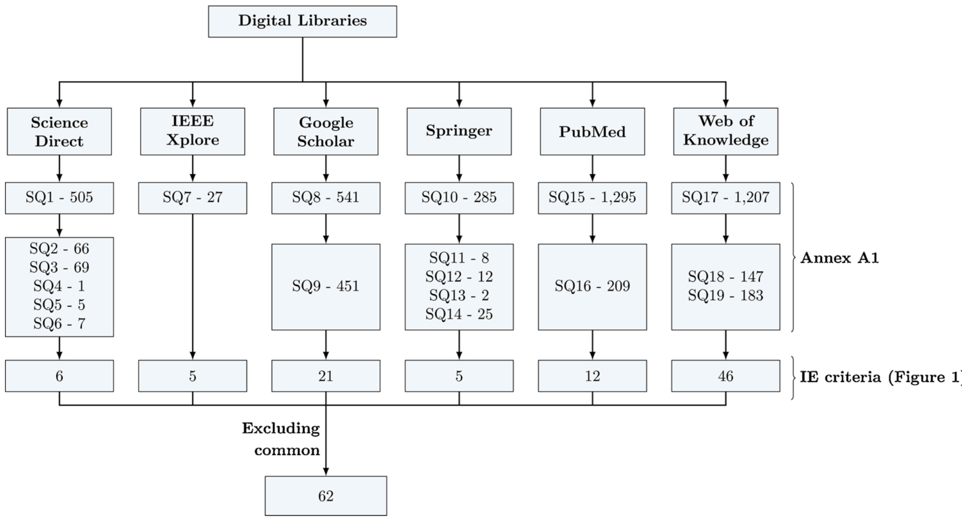

Fig. 2. Flow diagram of the digital libraries search and final selected papers after applying IE (Inclusion and Exclusion) criteria (search date 19/05/2019).

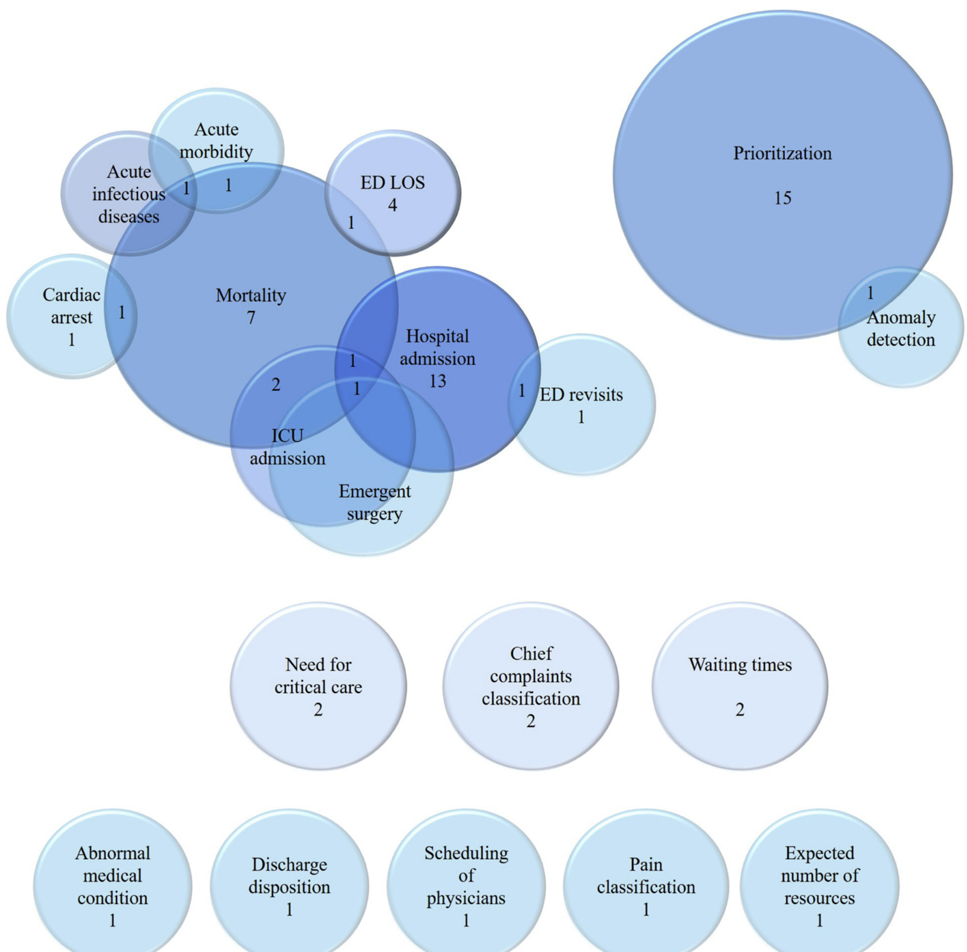

Fig. 3. Outcomes predicted in the selected papers with correspondent number of papers. ED - Emergency department, LOS - Length of Stay, ICU - Intensive Care Unit.

models to assist in the prioritization of patients, according to their acuity level at the triage, as presented in Fig. 3. Other objectives included prediction of hospital admission at discharge from the ED, ICU admission, ED LOS, mortality and need for critical care, abnormal medical condition, pain and chief complaints classification, acute morbidity and infectious diseases, cardiac arrest, ED revisits, discharge disposition, expected number of resources, scheduling of physicians and waiting times, from information available at the triage.

The triage system which was implemented in the ED in the majority of papers (34%) was Emergency severity index (ESI), followed by Patient Acuity Category Scale (8%) and Manchester Triage System (MTS) (6%). In the papers, triage priorities were whether used as input variables e.g. to predict patients' mortality [29], hospital admission [30], ED LOS [31], need for critical care [32] or consisted of the CDSS

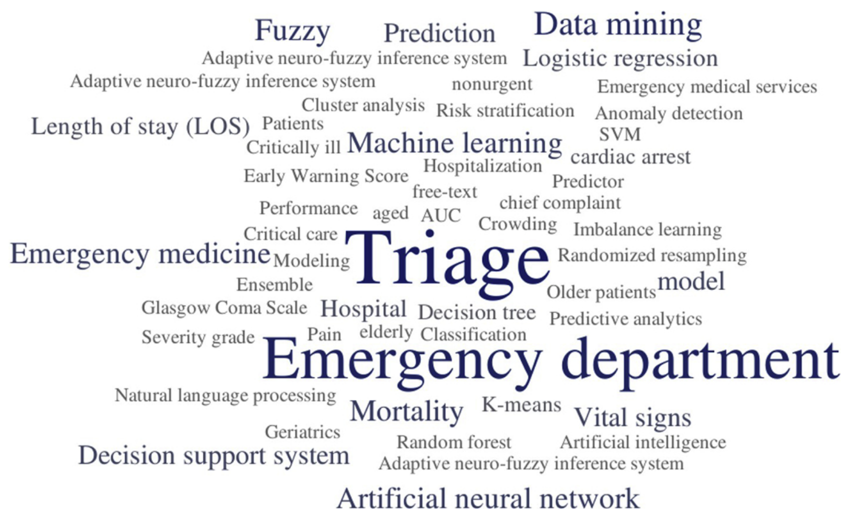

Fig. 4. Cloud of key-words of the papers selected in the review search.

output for triage classification [33].

The key words which were present in most of the selected papers were "triage" and "emergency department", which are emphasized in Fig. 4.

The variables most used in the papers for modelling were the ones presented in Fig. 5. The complete list of variables for each paper is presented in Table 4. Age (73%) and gender (66%) were used in more than half of the papers, heart rate in almost half of the papers (45%), followed by triage priorities (which were not used as input variables in papers where the objective was to predict the triage priority), chief complaints and SpO_2. Other vital signs, such as blood pressure, temperature and respiratory rate were also used by most papers to develop the models, as well as the patients' medical history, the arrival mode and time. The arrival mode comprised arrival from ambulance, referral, on foot and outpatient.

In about one third of the papers, different intelligent techniques were compared for predicting the outcome. The most used technique to develop the prediction models was logistic regression (LR), in 53% of the papers, as depicted in Fig. 6. About 40% of the papers selected only this technique for outcome prediction and in approximately 18% of the papers this technique was compared against other techniques, as depicted in Fig. 7. Classification and regression decision trees algorithms (CART) were used in 13% of the papers and the random forests classifier in 8%. Deep artificial neural networks (ANN) and support vector machines (SVM) were used in about 10% of the papers. Natural language processing (NLP) was used in papers where the unstructured free text was used as input variable to the model [34,35] or as the outcome [36,37].

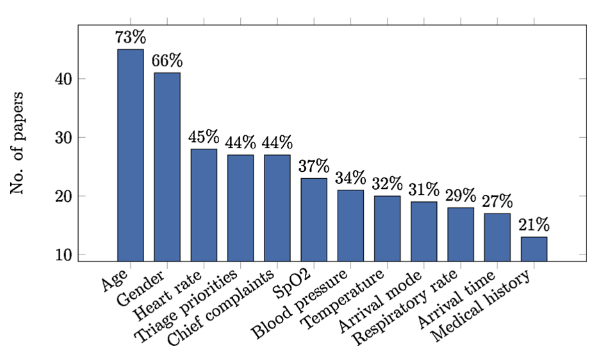

Fig. 5. Most used variables for the development of the triage models, with percentage value in the total number of selected papers.

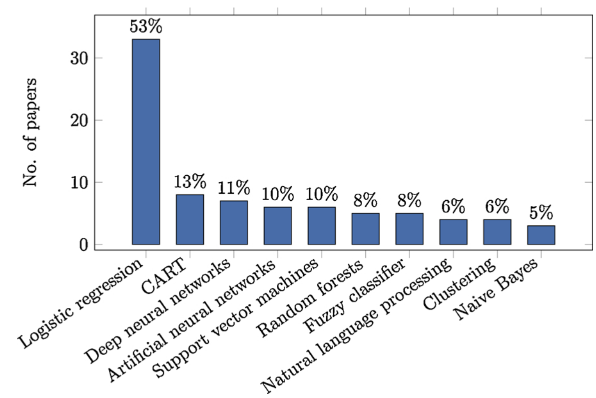

Fig. 6. Most used techniques for the development of the triage models, with percentage value in the total number of selected papers. CART - Classification And Regression Trees.

3.2. Prediction algorithms

In this subsection we briefly describe the most used machine learning techniques as well as those which presented higher performance compared to the other techniques, as presented in subsection 3.3.

3.2.1. Logistic regression

Logistic regression (LR) as a general statistical model was originally developed by Joseph Berkson [38]. LR is also designated "logit" model since it uses a "logit" link function which maps probabilities in the interval [0,1] to a real number. The logit of p_x, i.e. the probability of an event for certain covariate values x, is related to the covariates according to (1).

$$\text{logit}(p_x) = \log(\text{odds}_x) = \beta_0 + \beta_1 \times x \tag{1}$$

When compared to ANN, this model is less prone to over-fitting [39]. On the other hand, it has the drawback of having difficulties in handling non-linear problems and interactions between variables [39]. Nonetheless, LR is widely used in health research due to its easy interpretability, having been used in several studies [29–32,34,40–67].

3.2.2. Support vector machines

Support Vector Machines (SVM) [68] also denominated support vector networks [69] was used in several studies [35,46,57,70–72] and it creates optimal decision boundaries between data sets by solving a

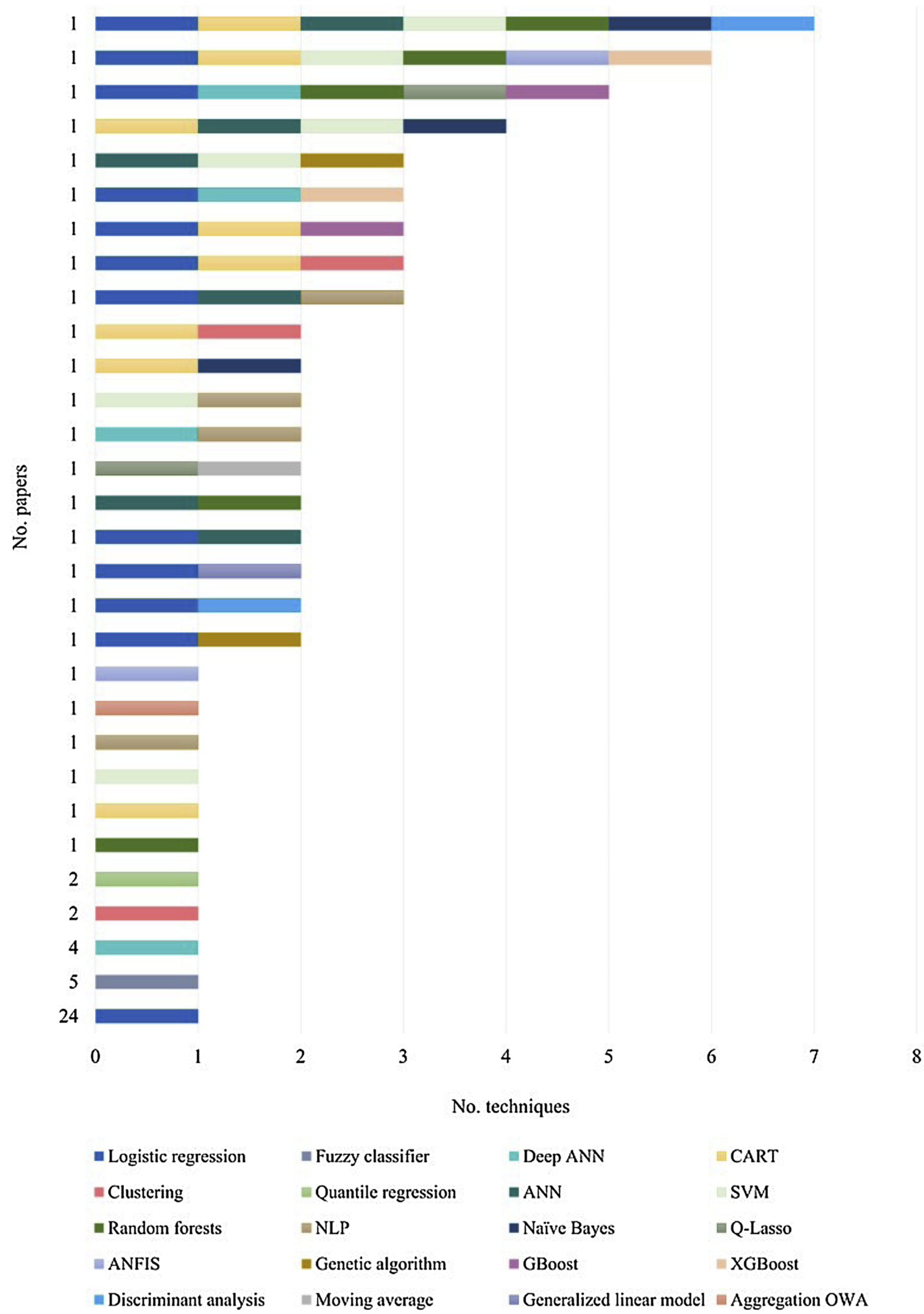

Fig. 7. Techniques used in the selected papers. ANN - Artificial neural networks, CART - Classification And Regression Trees, ANFIS - Adaptive Neuro Fuzzy Inference System, NLP - Natural Language Processing, SVM - Support Vector Machines, GBoost - gradient boosting, XGBoost - Extreme gradient boosting, OWA - ordered weighted average.

constrained quadratic optimization problem [73,74]. The disadvantage of SVM is that the classification result is purely dichotomous, and no probability of class membership is given. In other words, SVM attempt to draw a decision boundary that puts as many negative cases as possible on one side of the boundary and as many positive cases as possible on the other side. This makes it a non-probabilistic binary linear classifier. Another disadvantage of this model is that it is very sensitive to uncertainties [75], and prone to over-fitting in a high dimensional space [76]. However, it has a good generalization ability, being also robust for high dimensional data [77].

3.2.3. Naïve Bayes classifier

Naïve Bayes (NB) classifiers [78] are a family of simple "probabilistic classifiers" based on applying Bayes' theorem with strong (naive) independence assumptions between the features. NB is a conditional probability model, given an instance to be classified, represented by a vector $\mathbf{x} = (x_1,..., x_n)$, representing n independent variables, it assigns instance probabilities $p(C_k \mid x_1, .. ., x_n)$ for each k possible classes C_k. Using Bayes' theorem, the conditional probability can be decomposed as depicted in (2).

$$p(C_k|\mathbf{x}) = \frac{p(C_k)p(\mathbf{x}|C_k)}{p(\mathbf{x})} \tag{2}$$

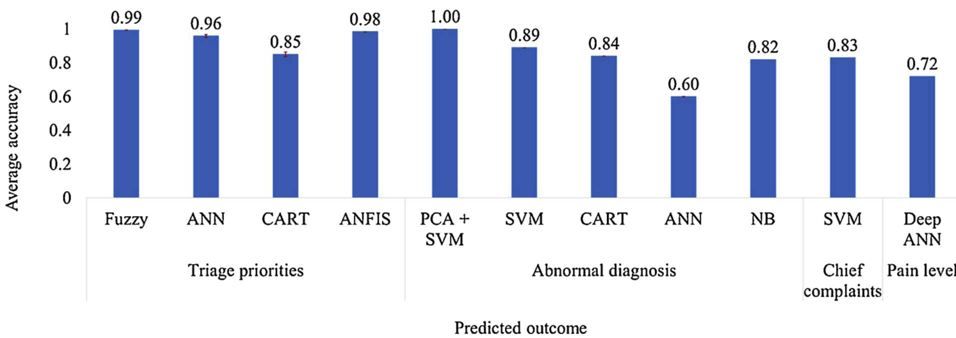

Fig. 8. Average performance in test achieved for each prediction goal, discriminated by technique, in the papers where the measure of accuracy was presented. ANN - Artificial neural networks, CART - Classification And Regression Trees, ANFIS - Adaptive Neuro Fuzzy Inference System, PCA - Principal Component Analysis, SVM - Support Vector Machines.

Using Bayesian probability terminology, (2) can be written as: posterior probability = (prior probability × likelihood)/ evidence.

This technique was used in the studies [57,71,79] and with appropriate pre-processing, it is competitive in the healthcare domain with other methods as e.g the case of prediction of abnormal diagnosis in Fig. 8, where it presents higher accuracy compared to ANN.

3.2.4. Fuzzy logic classifier

Fuzzy classifiers use fuzzy sets or fuzzy logic in the course of their training or operation [80]. Fuzzy logic represents a possibility logic model that uses reasoning to explain whether an event is about to happen [81]. This model was introduced by [82,83] and facilitates the process of vagueness treatment in a DSS by generating fuzzy rules using vague linguistic terms [84,85] instead of conventional rules to model decision boundaries in a more flexible way. However, it is difficult to estimate the membership functions [86].

A Fuzzy Inference System (FIS) uses fuzzy rules and fuzzy reasoning to perform its functions [87,88]. Mamdani inference system is usually used in clinical fuzzy systems [89] and its base structure includes four main components: a fuzzifier, which translates a crisp input into fuzzy values, i.e. soft labeling where a degree of membership to classes is generated; an inference engine, where in the case of Mamdani's inference, a fuzzy reasoning mechanism is applied to obtain a fuzzy output; a knowledge base, which contains a set of fuzzy rules, being the simplest ones of the type *if-then*, and a set of membership functions representing the fuzzy sets of linguistic variables and a defuzzifier, which translates the fuzzy output into crisp values. The fuzzy classifier was used in [90,91] and it has the advantage of interpretability compared with other techniques, such as ANN.

3.2.5. Artificial neural networks

Artificial neural networks (ANN) are suited for modelling non-linear problems, since they have the ability to learn without much in-depth understanding of the underlying system. The basic ANN consists of three layers, where the first one is the input layer, the second one the hidden layer and the third one the output layer, and each layer can have different numbers of neurons. There may be several inputs and outputs, as well as hidden layers. The hidden layer is connected in a feedforward or backpropagation manner to the input and output layers through different sets of weights. The most efficient and common architecture used in ANN is the feed forward ANN [92,93].

A deep neural network consists in an ANN with multiple layers between the input and output layers. Recurrent neural networks (RNN) [94] were used in [95] and they are a class of ANN where connections between nodes form a directed graph along a temporal sequence, which allows it to exhibit temporal dynamic behaviour. Long short-term memory (LSTM) was used in [96] and consists of an RNN architecture with feedback connections well suited to model time series data.

Multi-layer perceptron models do not make any assumptions on the distribution of data. Thus the non-linear form can be represented as in (3) where y is the output, w_i represents the synaptic weights, x_i denotes the covariates and I is the sigmoid activation function of the output. Weights in the model are assigned randomly at the beginning and then improved in an iterative training process.

$$y(x) = I\left(\sum_i w_i x_i - t\right) \tag{3}$$

ANN are robust to noisy data and have the ability to represent complex functions [97,98], however the use in clinical settings is limited because they are unable to explain decisions and lack transparency of data [97,99,100].

An adaptive neuro-fuzzy inference system (ANFIS) was used in [33,46] and it is an example of a hybrid intelligent system proposed in [101]. This system is capable of reasoning and learning in an uncertain and imprecise environment [102] and consists of a combination of two or more intelligent technologies, which allows it to overcome single intelligent technologies. Since the fuzzy system cannot learn or adapt by itself to the new environment and ANN are ambiguous to the user, when combining these two methods, the ANN becomes more transparent and the fuzzy system takes on the ability of learning.

3.2.6. Decision tree learning

Decision trees were used in [31,45,46,57,71,79,103,104] and have a tree-like hierarchy structure where each internal node has exactly two outgoing edges. The splits are selected using the towing criteria and the tree obtained is pruned by cost-complexity pruning. Each case follows appropriate branches until it reaches a terminal leaf node associated with a particular outcome. This learning algorithm automatically learns an optimal decision tree structure given a set of data. Classification trees are characterized by the target variable taking a discrete set of values while in regression trees the target variable takes continuous values. Trees formed may be unstable given their inefficiency for learning rules from incomplete data [105]. Compared with other machine learning methods, such as ANN, decision trees have the advantage that they are not black-box models, providing an easy interpretation of the classification system [97,39]. For this reason, decision trees are widely used in medicine.

Random forest [106] is an ensemble learning method that constructs a number of decision trees and it was used in a few studies [46,57,62,107,108]. For classification tasks, the random forests classifier outputs the class which is voted more times by the individual trees and for regression tasks it gives the mean prediction of the individual trees. The principle is that a group of trees – the“weak learners” can come together to form a “strong learner”. Random forests are able to correct the overfitting in training, which may be found in decision trees, by variance reduction [109]. This comes at the expense of some loss of interpretability, however a higher performance is achieved in the final model [39]. Moreover, this technique can easily handle outliers and is robust to inclusion of irrelevant features [109].

Gradient boosting (GBoost) used in [45,62] is a tree based ensemble technique which creates multiple weakly associated decision trees that are combined to provide the final prediction. The basic idea of boosting is to develop a new model in a gradient direction of the residuals to minimize the loss function generated at each iteration. Several models are developed and the weights are increased (boosted) if a model incorrectly classifies the observation. The extreme gradient boosting (XGBoost) algorithm is an efficient supervised learning algorithm which is a variant of the original GBoost method [110], having been very recently developed in [111]. XGBoost applies the second order Taylor expansion of the objective function, which has two parts, first the training loss and second the sum of the complexity of each tree, as given in (4). For a linear model given $\hat{y}_i = \sum_j \theta_j x_{ij}$ where $\hat{y}_i$ is the target variable, x_{ij}'s are the input variables and θ_j's are the coefficients of the model parameters, XGBoost iteratively applies greedy search to find the optimal model structure by adding a split to the existing tree structure at each iteration.

$$\mathrm{Obj}(\theta) = \sum_{i=1}^{n} L(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(y_k) \tag{4}$$

In (4) $L(y_i, \hat{y}_i)$ represents the loss function and $\Omega(y_k)$ the regularization term generated by each tree k, where K is the total number of trees.

3.3. Modeling performance assessment

The most used performance measures were accuracy and AUC, therefore we present a comparison between the algorithms discriminated by outcome for each performance measure in Figs. 8 and 9 , respectively. The outcomes presented in each figure were selected based on the criteria that in the majority of the studies for the respective outcome, either AUC or accuracy was presented. E.g. for studies of triage priorities classification only accuracy was assessed, and for prediction of mortality AUC was consistently assessed. We had the limitation of some of the studies not presenting any of these measures, therefore we present the average results considering a total of 41 papers. The performance results for all the papers is presented in annex, in Table 4.

For classification of triage priorities, Fuzzy modeling, ANN and ANFIS presented similar average performance. Fuzzy modeling presented the highest performance (99% accuracy), which was achieved in [91]. In the paper, a rule-based reasoning was applied for modeling the first to fourth decision points of the ESI algorithm and then combined with the fuzzy logic classifier. The variables used for modeling consisted in patients' age, gender and vital signs. The developed CDSS was able to reduce triage misdiagnosis and improve the triage outcomes.

For detection of abnormal medical conditions of patients at the triage, a combination of principal component analysis (PCA) and SVM led to an accuracy of 100%, which was achieved in [70]. Adding PCA to SVM improved accuracy from 89.2% to 100%. The variables used in the paper were patients' general appearance, chief complaints, medical history, vital signs, symptoms and signs, and physical assessment results. After periodic updates, the developed CDSS was able to improve the system without the influence of the subjective factor.

For prediction of chief complaints, an accuracy of 0.83 was achieved in [35] using SVM. In the paper, a system was developed by building an extended ontology of chief complaints and automatically predicting a patient's chief complaint, based on their vitals and the nurses' description of their state at arrival.

For binary pain intensity classification, an accuracy of 0.72 was achieved by a CDSS developed in [96] using LSTM recurrent neural networks. Audio-video recordings with indication of the location of the body pain, the pain level and a brief description on the type of pain felt were used. Vital signs and other clinically-related outcomes of on boarding emergency patients were also used.

In the papers where AUC was assessed, a higher number of techniques was used, as presented in Fig. 9. For the case of hospital admission, XGBoost presented the highest average performance of 0.89. In [47] an AUC of 0.92 was achieved with either XGBoost or deep learning, using the full list of variables: patients' age, gender, primary language, ethnicity, employment status, insurance status, marital status, and religion, the name of the hospital, arrival time, arrival mode, triage vital signs, ESI level, chief complaint, prior hospital and ED admissions, number of procedures and surgeries listed in the patient's record, medical history, medications, historical vitals and labs and number of orders for imaging.

For prediction of resource intensive patients, an average AUC of 0.88 was achieved in [95] using deep learning. Age and vital signs (structured data) and medical text (unstructured) data, including patient's chief complaint, past medical history, medication list, and nurse assessment were used for modeling. As for the need for critical care, the predictions ranged from 0.73 to 0.92 in [108] using random forests. The predictions of the CDSS demonstrated equivalent or improved identification of clinical patient outcomes compared with ESI. The variables used were age, gender, arrival mode (ambulance or walk-in), vital signs, chief complaint and relevant medical history.

For prediction of ED LOS using LR, in [65], an AUC of 0.80 was achieved using as input variables patients' age, usual accommodation, triage priority, arrival by ambulance, arrival overnight, imaging, laboratory investigations, overcrowding, time to be seen by doctor, ED visits with admission and access block relating to ED LOS more than 4 h. In [112], the performance for the prediction of this outcome was approximately equal, with an AUC of 0.79, using age, gender, triage priority, and final disposition decision from the ED.

For prediction of cardiac arrest, the higher average performance was 0.91 using LR. In [55], an AUC of 0.92 was achieved using patients' age, gender, comorbidities, functional status at presentation, mode of arrival, time of ED visit, triage priority at presentation, type of specialty, vital signs at different times, level of consciousness, need for supplement oxygen, need for ventilation assistance, use of a vasoactive agent (norepinephrine, epinephrine, or dopamine), use of an inotropic agent (dobutamine) and initial laboratory markers in the ED.

The acute morbidity prediction AUC was 0.82 in [57] using random

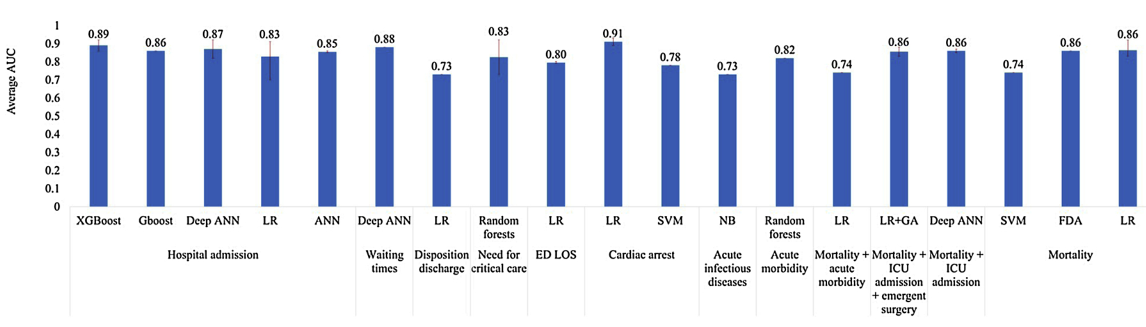

Fig. 9. Average performance in test achieved for each prediction goal, discriminated by technique, in the papers where the measure of Area Under the ROC Curve (AUC) was presented. LR - logistic regression, ANN - Artificial neural networks, GBoost - gradient boosting, XGBoost - Extreme gradient boosting, SVM - Support Vector Machines, NB - Naïve Bayes, GA - Genetic Algorithm, FDA - Flexible Discriminant Analysis.

forests and patients' age, gender, comorbidities, triage priority, chief complaints, vital signs, Glasgow Coma Scale score, medical history, physical examination and an electrocardiogram (ECG). In the paper, the results outperformed physicians' intuitive judgments.

Higher average mortality prediction performance was 0.86, either for the prediction of this outcome alone using LR, or for the combination with ICU admission using deep learning or with ICU admission and emergent surgery using LR and a genetic algorithm. The highest performance for mortality prediction was found in [44] with an AUC of 0.92 using LR and the variables patients' age, gender, time of admission, reason for admission, vital signs, Glasgow coma score, body peripheral perfusion and the presence of a threatened airway and information regarding interventions that had occurred before the time of ED arrival.

The outcomes of discharge disposition among medium acuity patients, mortality and acute morbidity, using LR, and acute infectious diseases, using Naïve Bayes, presented a lower average AUC of 0.73-0.74, compared to the other studies.

4. Discussion

4.1. Main contributions of the selected papers

The main contributions included the reduction of mortality rate, through the prediction of mortality [29,32,43,44,50,58,62–64,67,72], cardiac arrest [55,72], heart failure [56], acute morbidity and presence of acute infectious disease [57], development of acute renal failure, non-elective intubation, vasopressor requirement [60], as well as prediction of patients who should be admitted to the ICU [32,43,62] with the need for emergent surgery or catheterization [53]. Another important contribution consisted of the identification in due course of time of patients with significantly increased odds of hospital admission [30,34,37,41,42,45–49,54,59,62,113,114], leading to reduction of morbidity and mortality rate, improvement of patient pathways, prevention of readmissions and reduced costs for both the hospital and patients. Identification of high-risk patients [32], with significantly increased odds of hospital admission, could lead to improvement of resource allocation [40,42,54,108], potentially reducing ED overcrowding [42] and morbidity rate, which may be associated with delays in patients evaluation and treatment [40].

Other contributions included prediction of ED revisits [51,114], increased ED LOS [31,65,67,112,115], medium acuity patients' discharge disposition [61], improvement of ED bed management, average time to bed and rapid patients' discharge [30], the detection of overtriage and undertriage, performance of patients' retriage – thereby preventing adverse outcomes while waiting, reduction of unnecessary ED admissions, reduction of patients' waiting times [116,117] and improvement of resource allocation potentially reducing ED overcrowding. The reduction of the workload of triage health professionals was also an important contribution, as well as the improvement of scheduling of emergency physicians [104], or expected quantity of resources for critical patients [95], which resulted in an improvement of the overall quality of service for patients.

4.2. Limitations and future work

This subsection is structured according to the different topics presented by the authors of the selected papers when addressing limitations and future work.

4.2.1. Availability of data

Some authors [40,79,103,118,119] highlighted the importance of working with larger clinical datasets in order to apply intelligent techniques and extract knowledge. In [118], the authors asserted that a larger data set should be used to build a more accurate model. Regarding the variables used to develop the models, in addition to vital signs parameters, the authors in [103] acknowledged that complaints of patients should be considered in order to understand general principles of triage abnormal diagnosis. In [40], the authors stressed that variables such as time to surgery for surgical conditions or time to antibiotics for infections could have been included in the study and yielded better performance. In [42], only routine ED data collected at the time of triage were used for developing the model. The authors stated that there could be other important factors missed in the model, like the presenting symptoms, the vital signs, and functional or socio-economic status of patients. This fact may have limited the discrimination and validation power of the model. The authors stored the information on the patients' presenting symptoms and vital signs in electronic case notes in the format of free text. As future work, the authors offered a plan to extract this information into a structured data using text mining tools and to compare the model developed from the symptoms and vital signs to the model developed in the study. In [30], the authors argued that predictive models that incorporated more information that would become available as the healthcare process in the ED was progressing (clinical tests, preliminary clinical records, etc) would be likely to produce even better predictions and should be addressed in future research.

4.2.2. Geography

In [41], the authors recognized that the main limitation of the study was that, although it used data from different units, the hospitals were all in the same geographic region. This meant that they shared similar working practices, data recording methods, tertiary referral services and patient demographics. Ultimately, they concluded that, because it was an observational study, it was possible that there were unmeasured systematic biases particular to the region, therefore, no guarantee that the score's accuracy would hold elsewhere. Another limitation of the study was that, although the National Early Warning Score (NEWS) and the MTS are widely used in the UK, their inclusion would limit the score's use internationally. In [30], the authors suggested that, although the dataset was ample and represented a varied case mix, external validation, both in their country and in others, it would be desirable in order to assess wider applicability. In [32], it was concluded that the work was limited in that it was collected from one study site, which could constrain the generalizability of its results. In [119], it was suggested as future work the expansion of the study to include more hospitals, nurses, methods, or a more extensive or expansive scenario set.

4.2.3. Subjectivity of the system

Another limitation that was highlighted consisted of the subjectivity of the triage DSS, since these are reliant on the operator. In [30], the authors observed that with adequate training of the triage health professional, interoperator concordance was shown to be high in the case of MTS. In [79], the authors assessed the severity rankings of only two health professionals, bearing in mind the fact that they found the outcomes when learning from the consensus to be better than when learning from one health professional. Therefore, they recommended assessing a higher number of health professionals in future research and focus on the consensus classification of those health professionals. They asserted that this approach would enable the comparison between the classifiers' predictions from a majority of health professionals, which would lead to a much more robust measure. In [108], the authors stated that retrospective data are always subject to potential error in data entry. In their study, this was mitigated through data verification and processing that included chart review by a team of health professionals of their case study ED.

4.2.4. Methodologies and modelling techniques

Regarding methodologies and modelling techniques, some authors provided recommendations for future research. In [79], it was recommended that further studies should be performed using their proposed methodology. In [29], the authors found that further research would be required to identify an association of their DSS with access block (to inpatient beds) and patient mortality and also its possible use to predict the access block (to inpatient beds). In [120], the authors suggested that in the future their DSS should be compared to the rates of undertriage and overtriage. In [119], the authors suggested that their method could be implemented by itself and employed under peak or particularly stressful times if desired, or combined with a DSS for greater impact. In [41], it was highlighted that further work was needed to show that the DSS significantly outperformed triage nurses' predictions in a direct comparison, and to demonstrate the extent to which incorporation of the DSS into clinical practice actually improved care or use of resources.

4.2.5. Validation

For all the selected papers, real data from hospitals ED were used for model design. Half of the studies lacked the validation phase of their CDSS. In [108], it was described as a limitation that the study performed did not validate the tool's performance prospectively. In [44], the main limitations presented were related to the single-centre design and the lack of an external validation sample. The authors suggested that this reduced the generalizability of the model and created risk of overfitting. In the study bootstrapping was used for internal validation. In other papers [30,33,41,42], the authors performed an external validation where a dataset was used for model design and a smaller dataset was used for the validation phase, with a random data division. In e.g. [79], [107], [31] and [70], cross validation was performed. In [29], both internal and external validation were performed, first in a tertiary referral hospital and second in an urban community hospital.

5. Conclusions

In this paper, we presented a scoping review of ED triage decision support systems using machine learning. The main objectives of the papers selected from the review search consisted of the development of models for prioritization of patients or the prediction of hospital admission, ICU admission, ED LOS, mortality and need for critical care, abnormal medical condition, pain and chief complaints classification, acute morbidity and infectious diseases, cardiac arrest, ED revisits, discharge disposition, expected number of resources, scheduling of physicians and waiting times, from information available at the triage.

We concluded that the majority of the studies selected logistic regression as the modeling technique. When assessing the prediction models' performance, we found that the best Area Under the ROC Curve (AUC) results for emergency department length of stay, cardiac arrest and mortality prediction using this modeling technique were 0.80, 0.91 and 0.92, respectively. Deep neural networks also demonstrated high performance for prediction of resource intensive patients and hospital admission with an AUC of 0.88 and 0.92, respectively. Extreme gradient boosting presented equal performance for hospital admission. For classification of triage priorities, a rule-based reasoning combined with a fuzzy logic classifier yielded an accuracy of 0.99 and the prediction of abnormal diagnosis yielded and accuracy of 1 using principal component analysis with support vector machines. The variables included in these studies were patients' age, gender, vital signs, chief complaints, medical history, comorbidities, medication list, number of orders for imaging, laboratory variables, nurses assessment of patients' general physical appearance, emergency department arrival mode and time, time for medical observation and prior hospital and emergency department admissions. The triage priority level was also used as input variable in some of the studies.

We assessed how these systems have been assisting clinical decisions that resulted in improved quality of care at the ED triage. We found that the main contributions consisted of the identification in due course of time of patients with significantly increased odds of ICU or hospital admission and increased LOS, leading to reduction of morbidity and mortality rate, improvement of patient pathways, prevention of readmissions and reduced costs for both the hospital and patients. In the papers where the CDSS was validated in the ED, the authors found that there was an improvement of the health professionals' decision-making, which was more consistent and reliable. Some authors recognized that the main limitation was the lack of validation of the CDSS developed in a hospital. Furthermore, the CDSS should be designed to be flexible, so it can be implemented in any geography. Another limitation that was highlighted consisted in the subjectivity of the triage CDSS, since these are reliant on the operator. Thus, we stress that it is of foremost importance to assess the receptivity of health professionals to the use of the CDSS and that they receive adequate training.

We found that the majority of papers used real data in the design of the triage models. In these papers, the use of a set of variables for the development of the models was successful. However, the amount of information required by the triage models may be hampering the implementation of these systems. We stress that triage must be performed in about three minutes. This means that, aside from variables acquired and registered in the system at the ED admission, the triage models should include variables which are possible to acquire within this three minutes range (e.g. vital signs and chief complaints). As future work, some authors highlighted the importance of working with large clinical datasets in order to extract knowledge. It was also suggested that the studies should be multicenter, including more hospitals, nurses or methods. Furthermore, it was recommended to assess a higher number of health professionals when assessing the adoption of the decision support system. This approach enables the comparison between the classifiers' predictions from a majority of health professionals, leading to a more robust performance measure. It was also suggested that it is necessary to prove that the CDSS outperforms the health professionals' triage predictions, and to demonstrate the extent to which incorporation of the CDSS into clinical practice actually improves care or use of resources. Thus, we recommend that key performance measures of triage models are well defined such that decision makers are in a position to assess if an investment in such a system is viable.

In this scoping review we promoted an awareness of the relevance of intelligent CDSS in a complex and dynamic management environment, such as the hospital ED. This may help to create a bridge between intelligent CDSS and academicians and health professionals.

Conflicts of interest statement

No conflicts of interest.

Acknowledgements

This work was supported by the Portuguese Foundation for Science & Technology, through IDMEC, under LAETA, project UID/EMS/50022/2019 and LISBOA-01-0145-FEDER-031474 supported by Programa Operacional Regional de Lisboan by FEDER and FCT. The work of Marta Fernandes was supported by the PhD Scholarship PD/BD/114150/2016 from FCT.

Annexes

Table 2

Table 2
Results from the search queries (SQ) for each digital library (search date 19/05/2019).

Source	Query	Results
ScienceDirect	**SQ1**: Title, abstract or key-words: (triage) AND (”Emergency Department” OR “Emergency Room”) AND (“machine learning” OR modeling OR model OR classification OR predictive)	505
	SQ2: SQ1 refined by research and review articles and Publication Title “Annals of Emergency Medicine”	66
	SQ3: SQ1 refined by research and review articles and Publication Title “The American Journal of Emergency Medicine”	69
	SQ4: SQ1 refined by research and review articles and Publication Title “International Journal of Production Economics”	1
	SQ5: SQ1 refined by research and review articles and Publication Title “Expert System with Applications”	5
	SQ6: SQ1 refined by research and review articles and Publication Title “International Journal of Medical Informatics”	7
IEEE Xplore	**SQ7**: ((“Document Title”: triage) OR (“Abstract”: triage) OR (“Author Keywords”: triage)) AND ((“Document Title”: “Emergency Department”) OR (“Abstract”: “Emergency Department”) OR (“Author Keywords”:“Emergency Department”) OR (“Abstract”:“Emergency Room”) OR (“Document Title”:“Emergency Room”) OR (“Author Keywords”:“Emergency Room”)) AND ((“Document Title”:“machine learning”) OR (“Abstract”:“machine learning”) OR (“Author Keywords”:“machine learning”) OR (“Document Title”:modeling) OR (“Abstract”:modeling) OR (“Author Keywords”:modeling) OR (“Document Title”:model) OR (“Abstract”:model) OR (“Author Keywords”:model) OR (“Document Title”:classification) OR (“Abstract”:classification) OR (“Author Keywords”:classification) OR (“Document Title”:predictive) OR (“Abstract”:predictive) OR (“Author Keywords”:predictive)) with Full text &metadata	26
Google Scholar	**SQ8**:(intitle:triage) AND (intitle:“Emergency Department” OR intitle:“Emergency Room”) AND (“machine learning” OR modeling OR model OR classification OR predictive)	541
	SQ9: SQ8 -intitle:trauma -intitle:sepsis -intitle:children -intitle:pediatric -intitle:“chest pain”	451
Springer	**SQ10**: Where title contains: “triage” “emergency department”	285
	SQ11: SQ10 within Discipline: Engineering	8
	SQ12: SQ10 within Discipline: Computer Science	12
	SQ13: SQ10 within Discipline: Statistics	2
	SQ14: SQ10 within Discipline: Medicine &Public Health and Subdiscipline: Health Informatics	25
PubMed	**SQ15**: triage[tw] AND (“Emergency Department” [tw] OR “Emergency Room” [tw]) AND (“machine learning” [tw] OR modeling [tw] OR model [tw] OR classification [tw] OR predictive [tw])	1,295
	SQ16: SQ15 NOT (pediatric[Title] OR childhood[Title] OR paediatric[Title] OR child[Title] OR children[Title] OR pregnancy[Title] OR maternity[Title] OR mental[Title] OR depression[Title] OR dementia[Title] OR delirium[Title] OR suicide[Title] OR psychosocial[Title] OR psychiatry[Title] OR psychiatric[Title] OR resuscitation[Title] OR resuscitate[Title] OR dental[Title] OR dentist[Title] OR cancer[Title] OR sepsis[Title] OR influenza[Title] OR syncope[Title] OR stroke[Title] OR ischemic[Title] OR ischemia[Title] OR ischaemia[Title] OR thrombosis [Title] OR gastrointestinal[Title] OR HIV[Title] OR “abdominal pain”[Title] OR diabetic[Title] OR asthma[Title] OR cell[Title] OR pneumonia [Title] OR trauma[Title] OR traumatic[Title] OR chest[Title] OR myocardial[Title] OR chronic[Title] OR COPD[Title] OR pulmonary[Title] OR appendicitis[Title] OR poison[Title] OR poisoned[Title] OR hepatitis[Title] OR tomography[Title] OR ultrasound[Title] OR biomarkers[Title] OR bacteria[Title] OR bacterial[Title] OR virus[Title] OR overdose[Title] OR alcohol[Title] OR violence[Title] OR drug[Title] OR telehealth [Title] OR telephone[Title] OR catastrophe[Title] or epidemic[Title] OR epidemics[Title] OR epidemy[Title]) Filters activated: Free full text, English.	209
Web of Knowledge	**SQ17:** TS = (triage) AND (TS = (“Emergency Department”) OR TS = (“Emergency Departments”)) AND (TS = (“machine learning”) OR TS = (modeling) OR TS = (model) OR TS = (classification) OR TS = (predictive)))	1,207
	SQ18: SQ17 NOT (TI = pediatric OR TI = childhood OR TI = paediatric OR TI = child OR TI = children OR TI = pregnancy OR TI = maternity OR TI = mental OR TI = depression OR TI = dementia OR TI = delirium OR TI = suicide OR TI = psychosocial OR TI = psychiatry OR TI = psychiatric OR TI = resuscitation OR TI = resuscitate OR TI = dental OR TI = dentist OR TI = cancer OR TI = sepsis OR TI = influenza OR TI = syncope OR TI = stroke OR TI = ischemic OR TI = ischemia OR TI = ischaemia OR TI = thrombosis OR TI = gastrointestinal OR TI = HIV OR TI = ”abdominal pain” OR TI = diabetic OR TI = asthma OR TI = cell OR TI = pneumonia OR TI = trauma OR TI = traumatic OR TI = chest OR TI = myocardial OR TI = chronic OR TI = COPD OR TI = pulmonary OR TI = appendicitis OR TI = poison OR TI = poisoned OR TI = hepatitis OR TI = tomography OR TI = ultrasound OR TI = biomarkers OR TI = bacteria OR TI = bacterial OR TI = virus OR TI = overdose OR TI = alcohol OR TI = violence OR TI = drug OR TI = telehealth OR TI = telephone OR TI = catastrophe OR TI = epidemic OR TI = epidemics OR TI = epidemy) Refined by: LANGUAGE: (English) AND DOCUMENT TYPES: (ARTICLE OR PROCEEDINGS PAPER) AND WEB OF SCIENCE CATEGORIES: (ENGINEERING BIOMEDICAL OR HEALTH CARE SCIENCES SERVICES OR ENGINEERING MANUFACTURING OR MEDICAL INFORMATICS OR MULTIDISCIPLINARY SCIENCES OR COMPUTER SCIENCE INFORMATION SYSTEMS OR COMPUTER SCIENCE INTERDISCIPLINARY APPLICATIONS OR OPERATIONS RESEARCH MANAGEMENT SCIENCE OR COMPUTER SCIENCE ARTIFICIAL INTELLIGENCE OR ENGINEERING ELECTRICAL ELECTRONIC OR COMPUTER SCIENCE THEORY METHODS OR GERONTOLOGY OR MANAGEMENT OR ENGINEERING INDUSTRIAL)	147
	SQ19:SQ17 NOT (TI = pediatric OR TI = childhood OR TI = paediatric OR TI = child OR TI = children OR TI = pregnancy OR TI = maternity OR TI = mental OR TI = depression OR TI = dementia OR TI = delirium OR TI = suicide OR TI = psychosocial OR TI = psychiatry OR TI = psychiatric OR TI = resuscitation OR TI = resuscitate OR TI = dental OR TI = dentist OR TI = cancer OR TI = sepsis OR TI = influenza OR TI = syncope OR TI = stroke OR TI = ischemic OR TI = ischemia OR TI = ischaemia OR TI = thrombosis OR TI = gastrointestinal OR TI = HIV OR TI = ”abdominal pain” OR TI = diabetic OR TI = asthma OR TI = cell OR TI = pneumonia OR TI = trauma OR TI = traumatic OR TI = chest OR TI = myocardial OR TI = chronic OR TI = COPD OR TI = pulmonary OR TI = appendicitis OR TI = poison OR TI = poisoned OR TI = hepatitis OR TI = tomography OR TI = ultrasound OR TI = biomarkers OR TI = bacteria OR TI = bacterial OR TI = virus OR TI = overdose OR TI = alcohol OR TI = violence OR TI = drug OR TI = telehealth OR TI = telephone OR TI = catastrophe OR TI = epidemic OR TI = epidemics OR TI = epidemy) Refined by: LANGUAGE: (English) AND DOCUMENT TYPES: (ARTICLE OR PROCEEDINGS PAPER) AND WEB OF SCIENCE CATEGORIES: (EMERGENCY MEDICINE OR MEDICINE GENERAL INTERNAL) AND SOURCE TITLES: (ACADEMIC EMERGENCY MEDICINE OR AMERICAN JOURNAL OF EMERGENCY MEDICINE OR ANNALS OF EMERGENCY MEDICINE OR EUROPEAN JOURNAL OF EMERGENCY MEDICINE OR EMERGENCY MEDICINE JOURNAL OR JOURNAL OF EVALUATION IN CLINICAL PRACTICE OR EMERGENCY MEDICINE AUSTRALASIA OR SCANDINAVIAN JOURNAL OF TRAUMA RESUSCITATION EMERGENCY MEDICINE OR WESTERN JOURNAL OF EMERGENCY MEDICINE OR CANADIAN JOURNAL OF EMERGENCY MEDICINE OR MEDICINE OR ANNALS OF INTERNAL MEDICINE)	183

Table 3

Table 3
Sum up of the main findings of the papers selected in the review search.

Topic	Description
Main objectives	• To prioritize the patients, according to their acuity level, at the triage [33,66,70,79,90,91,103,107,108,118–124] and: predict abnormal diagnosis [103]; predict the need for critical care, an emergency procedure and inpatient hospitalization [108]; update patients' priority after their initial classification (retriage) [122] and to identify predictors of under- and over-triage [52]; • To predict exact number of resources and to identify resource intensive patients [95]; • To predict hospital admission with information available at the triage [30,34,37,40–42,45–49,54,59,62,108,113,114] and for a group of higher-risk patients, identified from among all those in a middle acuity triage category [40]; • To predict mortality with information available at the triage [29,32,43,44,50,53,56–58,62–64,67,72], ICU admission [32,43,53,62], acute morbidity and presence of acute infectious diseases [57], emergent surgery or catheterization [53]; • To predict ED LOS [31,65,67,112,115] from information available at the ED triage; • To predict medium acuity patients' discharge disposition [61]; • To predict patients' medical condition [71] or predictors of in-hospital adverse outcomes in ED patients with abnormal vital signs [60] from information available at the ED triage; • To predict patients' waiting time for medical observation [117,116]; • To improve scheduling of emergency physicians [104]; • To predict [35] or classify [36] patients' chief complaints; • To predict in-hospital cardiac-arrest with information available at the triage [55,63,72]; • To predict patients' revisits to the ED [51,114]; • To classify pain level during triage [96].
Intelligent techniques	• Random forests [46,57,62,107,108]; • Logistic regression (LR) [29–32,34,40–67] and general linear model [67]; • Q-Lasso [62,116] and moving average [116]; • Artificial neural networks (ANN) [30,33,34,37,57,66,71,113], deep learning [47,62,96], deep learning word attention model [95]; • Natural language processing (NLP) [34,35], computerized text-parsing algorithm for classification of free-text chief complaints/ coded chief complaint algorithm [36,37]; • Fuzzy classifier [90,91], fuzzy cognitive map structure [120,122–124]; • Aggregation methods, namely Borda-Kendall, aggregation through the estimation of utility intervals, aggregation using ordered weighted averaging (OWA) operator weights and a weight-determining model for rank aggregation [119]; • Adaptive neuro-fuzzy inference system (ANFIS) [33,46]; • Support Vector Machine (SVM) [35,46,57,70–72]; • Principal Component Analysis (PCA) and SVM for comparison with SVM and Back-propagation Neural Networks (BPNN). Support Vector Regression (SVR) and a Genetic Algorithm (GA) [70]; • LR and GA [53]; • Quantile regression [115,117]; • The Naïve Bayes (NB) algorithm [57,71,79]; • Decision trees for classification and regression (CART) algorithm [31,45,46,57,71,79,103,104]; • Linear, nonlinear and flexible discriminant analysis, partial least squares discriminant analysis, nearest shrunken centroids, K-nearest neighbors, J48 algorithm, PART rule [57]; • Multi-Group Discriminant Analysis [66]; • Gradient boosting (GBoost) [45,62] and extreme gradient boosting algorithm (XGBoost) [46,47]; • Clustering analysis: Ward's method [103], self-organizing-map [121], K-means [103,121,31], Fuzzy C-means and Fuzzy Subtractive algorithms [118].
Performance measures	• AUC [29,30,34,37,41,42,44–49,53–57,59–62,65,72,95,108,113]; • Specificity [32,37,45,60,62,71,72,79,107,113]; • Sensitivity [31,32,35,37,45,47,60,62,71,72,79,107,113]; • Accuracy (ACC) [33,35,45–47,66,70,71,91,95,96,104,107,118,120]; • Positive predictive value (PPV) [31,35,46,47,62,71,79]; • Negative predictive value (NPV) [47,62,72]; • F-score [31,35]; • Out-of-bag (OOB) error [107]; • Degrees of preference [125] and optimism levels [126] used in [119]; • Median difference between percentiles [115,117]; • Mean square error (MSE) [118], root MSE (RMSE) [33,116,118], %RMSE [33], absolute percentage error (APE) and mean APE (MAPE) [70]; • Importance weights for factor concepts (symptoms, medical history and vital signs) in each of the priority levels [124]; • Entropy change [79]; • Cohen's kappa (κ) used in [45] and in [103] to select best number of clusters and with expert decision, through graphical visualization of clusters gradient shades of color [121]; • Correlation laws from decision trees established based on confidence and support proportion [103]; • Hosmer-Lemeshow test [29,30] and Pearson's chi-squared (χ^2) test [30,41,42,52,112,114]; • Odds ratio (OR) [40,43,50,51,54,56,58,63,64,67,112,113].
Triage systems	• Emergency Severity Index (ESI) [32,34,35,37,46,47,49,52,53,57,61,62,95,108,113,115,119,120,122,124]; • Manchester triage sytem (MTS) [30,41,45,50]; • Australasian triage scale (ATS) [29,59]; • Canadian triage and acuity scale (CTAS) [40,56,114]; • Taiwan Triage and Acuity Scale (TTAS) which consists in a modification of CTAS adapted to EDs in Taiwan [54,96]; • Objective primary triage scale (OPTS) [33,107,118]; • Patient acuity category scale (PACS) [42,48,63,72,117]; • Simple Triage And Rapid Treatment (START) [71]; • A 3 level triage scale (immediate, urgent and standard, with target waiting times 0, 60 and 120 min, respectively) [90]; • French triage scale [51,112]; • Four-level Taiwan triage system [70,103,121]; • Rapid Emergency Triage and Treatment System-Adult (RETTS-A) in Sweden [58];

(continued on next page)

Table 3 (*continued*)

Topic	Description
Main contributions	• Adaptive Process Triage (ADAPT) in Sweden [64]; • No formal triage system, however, at the time there was an informal agreement regarding the semantics that determine a 5 level triage scale [79]; • Hillerøød Acute Process Triage (HAPT) system [43]. • Comparison of different intelligent techniques for patients' triage classification [30,33,40,45–47,70,71,79,103,107,118,119,121]; • Healthcare analysis of each patient, based on a health history report with ED admission information and triage results, and storage of patient's record in a database [90]; • Fast-track admission and outbound transfer planning [30,42], early warning system for the nursing personnel [30]; • Improvement of the prediction of LOS in the ED from triage information [31,65,67,112,115]; • Improved ED bed management [41], average TTB and rapid patients' discharge [30]; • Reduction of mortality rate, through the prediction of mortality from information available at the point of ED triage [29,32,43,44,50,58,62–64,67,72], cardiac arrest [55,72], heart failure [56], acute morbidity [57,60] and prediction of ICU admission [32,43,53,62]; • Performance of patients' retriage, thereby preventing adverse outcomes while waiting [122]; • Improved patient flow, decision support and control for demographics when comparing performance over time or between ED from different hospitals. Accurate estimate of the probability of hospital admission from triage information, which may improve patient pathways, prevent re-admissions and reduce costs [41,45–49]; • Reduction of unnecessary ED admissions for the elderly population [120]; • Improved pain level assessment to enhance the effectiveness of triage [96]; • Triage anomaly (overtriage and undertriage) detection [103,70,120,52]; • Reduction of operative costs and medical expenses for both the hospital and patients [121]; • Reduction of human error [107] and improvement of service quality for patients [121]; • Improvement of the prediction for the need of hospitalization and critical care [108]; • Prediction of individual patients' waiting time [117,116]; • The indication that the best time frame to evaluate the revisits rate after an ED visit is 30 days and the suggestion to use unscheduled revisits to the ED as an indicator of quality [51,114]; • Improvement of scheduling of emergency physicians [104] and expected number of resources for critical patients [95]; • Identification of high-risk patients [32], with significantly increased odds of hospital admission, leading to improvement of resource allocation [40,42,54], potentially reducing ED overcrowding [42] and morbidity rate, which may be associated with delays in patients evaluation and treatment [40].

Table 4

Table 4
Results by outcome according to the objectives of the papers selected in the review search. Where omitted, the performance values indicated were obtained in the test set.

Objectives	Author (Year; Source)	Independent Variables	Results by outcome (Algorithm: performance)
Predict in-hospital cardiac arrest	W. Srivilaithon et al. [55] (2019; Emergency Medicine Australasia)	Age, gender, comorbidities, functional status at presentation, mode of arrival, time of ED visit, triage level at presentation, type of specialty, physiological parameters at different times (respiratory rate, oxygen saturation, temperature, blood pressure, and heart rate), level of consciousness, need for supplement oxygen, need for ventilation assistance, use of a vasoactive agent (norepinephrine, epinephrine, or dopamine), use of an inotropic agent (dobutamine) and initial laboratory markers in the ED.	LR using all predictors AUC 0.91 (95% CI 0.89-0.93) higher than NEWS alone model AUC 0.78 (95% CI 0.74-0.81).
Predict in-hospital cardiac arrest and mortality	M. Ong et al. [72] (2019; Critical Care)	Age, gender, medical history including ischemic heart disease, diabetes mellitus and chronic renal failure, heart rate, blood pressure, respiratory rate, oxygen saturation (SpO2), Glasgow Coma Scale (GCS), etiology, and medication history.	Prediction of cardiac arrest SVM AUC: 0.781, compared with 0.680 for MEWS (difference in AUC: 0.101, 95% CI: 0.006 to 0.197). A cutoff machine learning score ≥ 60 predicted cardiac arrest with a sensitivity of 84.1%, specificity of 72.3% and NPV of 98.8%. Prediction of in-hospital mortality SVM AUC: 0.741, compared with 0.693 for MEWS (difference in AUC: 0.048, 95% CI: -0.023 to 0.119). A cutoff MEWS ≥ 3 predicted mortality with a sensitivity of 74.4%, specificity of 54.2% and NPV of 97.8%.
Predict mortality, acute morbidity and presence of acute infectious disease	M. Jenny et al. [57] (2015; Academic Emergency Medicine)	Age, gender, comorbidities, priority levels, chief complaints, vital signs, Glasgow Coma Scale score, medical history, physical examination and an electrocardiogram (ECG). In addition, we obtained a measure of the physician's first overall impression of each patient, the Gestalt-like impression of "how ill the patient looks".	Machine learning models (random forests, LR, ANN, SVM, NB, CART): predictability of the target outcomes ranged between AUC of 0.71 and 0.82. These results outperformed physicians' intuitive judgements (AUC = 0.67 for mortality, 0.65 for morbidity, and 0.60 for infectious disease).
Predict heart failure mortality			

(continued on next page)

Table 4 (*continued*)

Objectives	Author (Year; Source)	Independent Variables	Results by outcome (Algorithm: performance)
	D. Lee et al. [56] (2012; Annals of Internal Medicine)	Age, gender, transport to ED by emergency medical services, origin from nursing home or long-term care facility, initial vital signs (systolic blood pressure, heart rate, respiratory rate, and oxygen saturation), symptoms and comorbidities, laboratory results, atrial fibrillation, electrocardiogram, and pre-ED medications obtained by medical staff and recorded in the patient's chart and priority level.	LR AUC: 0.805 for the derivation data set (bootstrap-corrected, 0.811) and 0.826 for validation data set.
Predict adverse outcomes (development of acute renal failure, non-elective intubation, vasopressor requirement, or mortality)	D. Henning et al. [60] (2015; Western Journal of Emergency Medicine)	Age, gender, vital signs, past medical history, length of stay, laboratory values.	LR AUC, sensitivity and specificity: 0.74, 0.70 and 0.63, respectively.
Predict critical outcomes (mortality, ICU admission, emergent surgery or catheterization)	S. Barnes et al. [53] (2018; Journal of Healthcare Engineering)	Age, gender, arrival mode of the patient, vital signs (heart rate, respiratory rate, temperature, blood pressure, and oxygen saturation) and the chief complaints.	LR using GA flattened approach AUC: a large urban Academic medical center (ACAD) - 0.8431, a medium-sized community hospital (COMM) - 0.8361, international hospitals in Brazil (BRAZIL) - 0.8261, the United Arab Emirates (UAE) - 0.8820, and the nationally representative National Hospital Ambulatory Medical Care Survey (NHAMCS) - 0.8429. LR using GA hierarchical approach AUC: ACAD - 0.8433,COMM - 0.8364, BRAZIL - 0.8260, UAE - 0.8819, NHAMCS - 0.8436.
Predict mortality	M. Ljunggren et al. [58] (2016; Scandinavian journal of trauma, resuscitation and emergency medicine)	Age, gender, time, date, admittance to in-hospital care and, if so, to which clinic, were recorded during the ED visit. Presenting symptoms, SpO2 (%), respiratory rate, heart rate, systolic blood pressure, diastolic blood pressure, temperature and level of consciousness according to the AVPU scale, the triage priority level, comorbidities. Data regarding the presence of a threatened airway, oxygen use, and whether the heart was regular or irregular was incorporated in the vital sign triage prioritisation assessment.	LR highest OR: unresponsive vs alert patients (OR 31.0, CI 16.9 to 56.8), patients with more than 80 years old with less than 50 years old (OR 35.9, CI 10.7 to 120.2) and patients with respiratory rates $<$ 8/min to 8 to 25/min (OR 18.1, CI 2.1 to 155.5).
	M. Coslovsky et al. [44] (2015; Intensive Care Medicine Journal)	Age, gender, time of admission, cause of admission, respiratory rate, oxygen saturation, systolic and diastolic blood pressure, heart rate, Glasgow coma score, body temperature and peripheral perfusion and the presence of a threatened airway and information regarding interventions that had occurred before the time of ED arrival.	LR: AUC 0.92 (95% CI, 0.916-0.927).
	D. J. Teubner et al. [29] (2015; Emergency Medicine Australasia)	Age at presentation, gender, ATS level, transport to the ED by ambulance, referral to the ED by a physician and triage complaint category - a total of 108 variables were assessed as predictors of inpatient mortality.	LR: AUC for the train set was 0.859 (95% CI 0.856-0.865), for the internal validation set was 0.848 (95% CI 0.840-0.856) and for the external validation set was 0.837 (95% CI 0.823-0.851);
	T. Djärv et al. [64] (2015; European Journal of Emergency Medicine)	Age, gender, and disease burden (number of medications and chronic diseases), mode of arrival to the ED and triage priority.	LR: patients with decreased general condition at the ED were admitted for in-hospital care, and they had a four-fold risk of suffering an in-hospital death [OR 4.74 (95% CI 3.88-5.78)] compared with patients presenting with other presenting complaints.
	W. Hong et al. [63] (2013; European Journal of Emergency Medicine)	Heart rate, systolic blood pressure, diastolic blood pressure, respiratory rate, oxygen saturation, and the Glasgow Coma Scale.	Vital signs with age LR sensitivity: cardiac arrest 11.54%, death 22.73%, ICU 12.50% Vital signs with age LR specificity: cardiac arrest 99.28%, death 97.22%, ICU 93.80%.
	P. Plunkett et al. [50] (2011; European Journal of Emergency Medicine)	Gender, major disease by priority level, Charlson's comorbidity index, ICU admission, blood transfusion, troponin elevation, door-to-team and team to- ward time.	After adjustment for all outcome predictors, including comorbidity and illness severity, the results for both independent predictors of death within 30 days were: door-to-team times OR: of 1.13 (95% CI 1.07-1.18) and team-to-ward times OR: 1.07 (95% CI 1.02-1.13).

(*continued on next page*)

Table 4 (*continued*)

Objectives	Author (Year; Source)	Independent Variables	Results by outcome (Algorithm: performance)
Predict mortality, ICU admission and hospital admission	Y. Raita et al. [62] (2019; Critical Care)	Age, gender, triage vital signs, chief complaints and comorbidities.	For critical care outcome prediction, AUC: deep ANN 0.86 [95%CI 0.85-0.87] vs LR reference model 0.74 [95%CI 0.72-0.75]. For the hospitalization outcome prediction, AUC: deep ANN 0.82 [95%CI 0.82-0.83] vs LR reference model 0.69 [95%CI 0.68-0.69].
Predict mortality and ICU admission	M. LaMantia et al. [32] (2013; Western Journal of Emergency Medicine)	Age, gender, ESI score, chief complaint and vital signs. Vital signs consisted of systolic BP, heart rate, respiratory rate, temperature and oxygen saturation.	LR: sensitivity and specificity of 73% (95% CI 66-81) and 50% (95% CI 48-52), respectively (positive likelihood ratio 1.47 (95% CI 1.30-1.60) and negative likelihood ratio 0.54 (95% CI 0.30-0.60).
	C. Barfod et al. [43] (2012; Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine)	Age, gender, time of the day for contact (morning, evening, night), weekday (week-end versus week-day), vital signs (SpO2, respiratory rate, systolic BP, heart rate), chief complaint and GCS.	The vital signs, the chief complaint and triage priority were all significantly associated with ICU admission and in-hospital mortality, the odds increasing with the urgency of the triage priority.
Predict mortality and ED LOS	Y. Ro et al. [67] (2015; Emergency Medicine Australasia)	Age, gender, vital signs, mental status, method of transportation, reason for visiting the ED (medical illness or trauma), priority level and number of all occupant patients in the ED at the time for every h/min when each patient visited the ED.	LR adjusted OR (95% CI) for the triage-based resource allocation and clinical treatment (TRACT) protocol on ED mortality: 0.69 (0.54-0.88) for total patients, 0.42 (0.30-0.59) for ESI 1, 1.04 (0.66-1.65) for ESI 2 and 1.45 (0.76-2.75) for ESI 3 group. LR adjusted OR (95% CI) for the TRACT protocol ED LOS: -88.1 (-96.9 ∼ -79.2) min for all patients, -44.9 (-72.0 ∼ -17.9) min for ESI level 2 and -104.3 (-114.7 ∼ -94.0) min for ESI level 3.
Predict expected number of resources	D. Gligorijevic et al. [95] (2018; Proceedings of the 2018 SIAM International Conference on Data Mining)	Age, hearth rate, blood pressure, temperature (structured data) and medical text (unstructured) data, including patient's chief complaint, past medical history, medication list, and nurse assessment.	Deep learning word attention model AUC: 0.88 in identifying resource intensive patients (binary classification); Deep learning word attention model ACC: 44% for predicting exact category of number of resources (multi-class classification task).
Predict patients' waiting times	E. Ang et al. [116] (2015; Manufacturing &Service Operations Management)	The number of patients waiting in the ED to start treatment, the number of providers in the ED, the rate at which a provider treats low-acuity patients and the total processing rate for low-acuity patients.	In the triage room, Q-Lasso achieved a 30% lower MSE in predicting the residual time from triage to treatment than would have occurred with another method – the two-hour-window rolling average. Q-Lasso MSE: 998.6 minutes with a standard error of 24.0 minutes vs moving average MSE: 1,429.4 minutes with a standard error of 35.9 minutes.
	Y. Sun et al. [117] (2012; Annals of Emergency Medicine)	Patient queue sizes, flow rates and patient priority level.	Quantile regression in retrospective validation: median absolute prediction error was 11.9 minutes for patient acuity priority 2 (interquantile range (IQR) 5.9 to 22.1 minutes) and 15.7 minutes for priority 3 (IQR 7.5 to 30.1 minutes); In prospective validation: median absolute prediction error was 9.2 minutes for patient acuity priority 2 and 12.9 minutes (IQR 6.5 to 22.5 minutes) for priority 3.
Predict ED revisits	L. Pereira et al. [51] (2015; PloS one)	Age, gender, triage priority, care pathways, diagnostic categories based on International Classification of Diseases, Ninth Revision (ICD-9) codes, final disposition, the time of arrival and departure from the ED Observation Unit (OU) as well as the departure date for hospitalized patients.	LR indicated that the ED final disposition decision (transfer; medical and surgical wards (MSW); non-admission) was a significant 30-day predictor of revisit to the ED (OR: 1.52 (95% CI 1.42-1.56); 2.32 (2.12-2.38); $p < 0.0001$]; and that care pathways were significant 90-day predictors (ED→OU→MSW; ED→OU→transfer; ED transfer; ED→MSW; ED→OU→non-admitted; ED non-admitted) (OR 1.11 (95% CI 1.08-1.12); 1.23 (1.19-1.24); 1.35 (1.31-1.36); 1.50 (1.44-1.51); 1.66 (1.64-1.67); $p < 0.001$).
Predict hospital admission and ED revisits	A.Hendin et al. [114] (2018; Canadian Journal of Emergency Medicine)	Age, gender, tests and services involved in ED, and disposition. Return ED visit and hospital admission rates at 14 days were tracked.	Older patients were significantly more likely than younger controls to be admitted on index visit (5.0% vs 0.3% admit rate, $p = 0.016$). They had a trend towards increased re-presentation rates within 14 days (13.7% vs 8.7% control, $p = 0.11$) and were more likely to be admitted on re-presentation

(*continued on next page*)

Table 4 (*continued*)

Objectives	Author (Year; Source)	Independent Variables	Results by outcome (Algorithm: performance)
			(4.0% vs 0.7%, p = 0.045). In sub-group analysis, very elderly patients (85 years and up, n = 79) were more likely to be admitted (8.9%, p = 0.003).
Predict medium acuity patients' discharge disposition	J. Riordan et al. [61] (2017; The Journal of emergency medicine)	Age, gender, ED disposition, arrival mode, temperature, systolic and diastolic blood pressure, heart rate, respiratory rate, oxygen saturation, and pain level.	LR AUC: 0.73.
Improve scheduling of emergency physicians	C. Yang et al. [104] (2009; Expert Systems with Applications)	Date of diagnosis, triage priority, department, number of nurses and physicians and their daily shifts, day of the week and national holiday or not.	Decision tree ACC of shift anticipation improved from 22% to 50%.
Predict hospital admission	O. Araz et al. [46] (2019; OPENAIRE)	Age, gender, arrival mode, triage priority, day of visit, flu season, shift, and Influenza-Like-Illness case indicator (using ICD-9 codes).	XGBoost: AUC ranged from 0.83 to 0.86.
	B. Graham et al. [45] (2018; IEEE Access)	Hospital site, date and time of attendance, age, gender, arrival mode, care group, MTS priority and whether the patient had a previous admission to the hospital within the last week, month, or year.	GBoost: ACC 80.31%, AUC 0.86; DT: ACC 80.06%, AUC 0.82; LR: ACC 79.94%, AUC 0.85.
	C. Parker et al. [48] (2018; The American journal of emergency medicine)	Age, gender, ethnicity, proximity of patient's home postal code to the study site, day of week, shift time of presentation, mode of arrival, triage category, fever status and number of ED visits within the previous year.	LR: AUC 0.82 (95% CI 0.82-0.83).
	W. Hong et al. [47] (2018; PloS one)	Age, gender, primary language, ethnicity, employment status, insurance status, marital status, and religion. Name of the hospital, arrival time, arrival mode, triage vital signs (systolic and diastolic blood pressure, heart rate, respiratory rate, oxygen saturation, presence of oxygen device and temperature), and ESI level. Chief complaint, prior hospital and ED admissions, number of procedures and surgeries listed in the patient's record, medical history (ICD-9 codes), medications, historical vitals and labs and number of orders for imaging.	Using triage information in train: LR: AUC 0.87 (95% CI 0.86-0.87); XGBoost: AUC 0.87 (95% CI 0.87-0.88); Deep ANN: AUC 0.87 (95% CI 0.87-0.88). Using patient history in train: LR: AUC 0.86 (95% CI 0.86-0.87); XGBoost: AUC 0.87 (95% CI 0.87-0.87); Deep ANN: AUC 0.87 (95% CI 0.87-0.88). Using the full set of variables in train: LR: AUC 0.91 (95% CI 0.91-0.91); XGBoost: AUC 0.92 (95% CI 0.92-0.93); Deep ANN: AUC 0.92 (95% CI 0.92-0.92).
	X. Zhang et al. [34] (2017; Methods of information in medicine)	Unstructured variables: three reasons for visit and cause of injury. Structured information: age, gender, race, ethnicity, type of residence, source of payment, whether or not arriving via ambulance, arrival day and time, initial vital signs: body temperature, heart rate, respiratory rate, blood pressure, heart oximetry, whether or not the patient arrived on oxygen. Triage variables: priority level, pain scale (0-10), whether or not the patient had used the ED within the past 72 hours, the episode of care (initial vs. follow-up visit to the ED) and comorbidities and injury.	Models including only structured variables: LR AUC 0.824(95% CI 0.818-0.830), ANN AUC 0.823 (95% CI 0.817-0.829). Models including only free-text information: LR AUC of 0.742 (95% CI 0.731-0.753), ANN 0.753 (95% CI 0.742-0.764). Models with both structured variables and free text variables: LR AUC 0.846 (95% CI 0.839-0.853), ANN AUC 0.844 (95% CI 0.836-0.852).
	C. Ng et al. [54] (2016; Medicine)	Age, gender, primary and secondary diagnoses, clinical parameters including blood pressure, heart rate, body temperature, and chief complaint category in TTAS	LR AUC: 0.70;
	A. Zlotnik et al. [30] (2016; CIN: Computers, Informatics, Nursing)	Repeated visits, age, gender, patient insurance status and main residence, patient visit sources, broad visit causes, ambulance arrivals, MTS score and MTS chief complaints, which were combined in five groups of increasing risk of admission.	LR: AUC 0.86 (95% CI, 0.8508-0.8583); ANN: AUC 0.86 (95% CI, 0.8540-0.8610).
	N. Handly et al. [37] (2015; European Journal of Emergency Medicine)	Age, gender, race, time, and day of arrival, initial priority level, chief complaint and whether the patient was admitted to the hospital or not on each ED visit.	ANN AUC in the validation cohort for the derived models with and without the coded chief complaint algorithm: 0.840 (95% CI: 0.838-0.842) and 0.860 (95% CI: 0.858-0.862), respectively.
			LR: AUC 0.88 (95% CI 0.8752-0.8796).

(*continued on next page*)

Table 4 (*continued*)

Objectives	Author (Year; Source)	Independent Variables	Results by outcome (Algorithm: performance)
	A. Cameron et al. [41] (2014; Emergency Medicine Journal)	Age, gender, means of transportation to the ED, the time of the day and of the week of attendance, the referral source, the MTS category, the NEWS, if the patient lived alone or not and previous admissions.	
	S. Kim et al. [59] (2014; Emergency Medicine Australasia)	Age, gender, day of the week, time of the day, triage priority, whether the patient arrived by ambulance or not, and whether the patient was referred by a local medical officer.	LR AUC: Model 0 – characteristics of each patient excluding any laboratory test results 0.797 (95% CI 0.795, 0.800) Model 1 - triage nurses' prediction 0.749 (95% CI 0.746, 0.751) Model 2 – Model 0 + Model 1 0.817 (95% CI 0.815, 0.819) Model 3 – Model 2 + presence/absence of laboratory test 0.835 (95% CI 0.833, 0.837) Model 4 – Model 3 + laboratory test results 0. 768 (95% CI 0.764, 0.772);
	N. Handly et al. [113] (2013; Annals of Emergency Medicine)	Age, gender, race, time and day of arrival, priority level, chief complaint and admission/ discharge data on each visit.	ANN AUC in the validation cohort for the prediction model without coded chief complaint data: 0.840 95%CI (0.838-0.842). ANN AUC in the validation cohort for the prediction model with coded chief complaint data: 0.860 95%CI (0.858-0.862).
	Y. Sun et al. [42] (2011; Academic Emergency Medicine)	Age, gender, ethnic group, ED visit or hospital admission in the preceding 3 months, arrival mode, patient priority, and coexisting chronic diseases.	LR: AUC 0.85 (95% CI 0.847-0.851).
	M. LaMantia et al. [49] (2010; Academic emergency medicine)	Age, triage priority, heart rate, diastolic blood pressure, and chief complaint.	LR: AUC 0.73;
Predict abnormal medical condition of patients	D. Olivia et al. [71] (2018; International Conference on Applications and Techniques in Information Security)	Body temperature, blood pressure (diastolic and systolic), blood oxygen level, respiration rate, and heart rate.	ANN: ACC 0.60, Sensitivity 0.87, Specificity 0.18, PPV 0.61; NB: ACC 0.82, Sensitivity 0.84, Specificity 0.80, PPV 0.86; DT, SVM: ACC 0.84, Sensitivity 0.87, Specificity 0.81, PPV 0.87.
Classify pain level	F. Tsai et al. [96] (2017; Seventh International Conference on Affective Computing and Intelligent Interaction (ACII))	Audio-video recordings with indication of the location of the body pain, the pain level and a brief description on the type of pain felt. Physiological (heart rate, systolic and diastolic blood pressure) vital sign data, and other clinically-related outcomes of onboarding emergency patients.	LSTM: ACC 72.3% and 54.2% in binary and three-class pain intensity classification, respectively.
Predict the need for critical care and of high-risk patients	S. Levin et al. [108] (2017; Annals of Emergency Medicine)	Age, gender, arrival mode (ambulance or walk-in), vital signs (temperature, heart rate, respiratory rate, systolic blood pressure, and oxygen saturation), primary chief complaint, and relevant medical history.	Random forest model: AUC ranged from 0.73 to 0.92.
	J. P. Ruger et al. [40] (2007; The American Journal of Emergency Medicine)	Age and gender, primary complaint variables and a diagnosis related group (DRG) severity index. Other variables included medical diagnosis (DRG diagnosis), day and time of the week, month of the year, arrival mode, and payment method.	In the LR analysis, acuity C patients of 65 years or older were 2.5 times more likely to be admitted than those younger than 65 years old (*p*-value $<$ 0.05).
Predict chief complaints	Y. Jernite et al. [35] (2013; NIPS 2013 Workshop on Machine Learning for Clinical Data Analysis and Healthcare)	Chief complaints, vital signs and the nurses' description of the patients' state at arrival.	Multiclass SVM for the best-10 ACC: 0.825. The best-10 ACC measured how often the list of 10 most likely predicted labels actually contained all of the true chief complaints.
	A.Thompson et al. [36] (2006; Academic emergency medicine)	Age, gender, admission rate, frequency of patients presenting with coded chief complaints, percentage of free-text complaints not categorizable by the proposed computerized text-parsing algorithm for classification of free-text chief complaints.	In the derivation sample, the text-parsing algorithm classified 87.5% of 45,329 ED visits with non null free-text chief complaints into 1 of 194 coded chief complaints. The text-parsing algorithm successfully classified 87.3% of the free-text chief complaints in a validation sample.
Prioritize patients	M. Soufi et al. [91] (2018; International journal of medical informatics);	Age, gender and vital signs (heart rate, SPO2, respiratory rate).	A combination of the Rule-Based Reasoning (RBR) and Fuzzy Logic Classifier (FLC) ACC: 99.44%.
	V. Georgopoulos and C. Stylios [123] (2017; 2017 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS))	Variables used in [122].	The Fuzzy Cognitive Map indicated that patients with a fever and infections could adversely progress over time, which meant their priority should be upgraded. Therefore,

(*continued on next page*)

Table 4 (*continued*)

Objectives	**Author** (Year; Source)	**Independent Variables**	**Results by outcome** (Algorithm: performance)
			the Fuzzy Cognitive Map was designed to make sure that patients would not wait for hours not receiving the necessary care.
	J. Hinson et al. [52] (2018; International Journal of Emergency Medicine)	Age, gender, temperature, heart rate, respiratory rate, systolic pressure, oxygen saturation, patient-reported pain score, chief complaint category, and arrival time.	Under-triaged patients to ESI level 3 had a significantly increased prevalence of admission and critical outcomes as compared to those appropriately triaged to ESI level 3 χ^2 = 502.06, df = 1, p value < 0.001 and χ^2 = 184.91, df = 1, p value < 0.001, respectively). Similarly, patients who were under-triaged to ESI levels (4 or 5) on arrival had a significantly increased prevalence of admission and critical outcomes as compared to patients appropriately triaged to the same ESI levels (χ^2 = 1033.60, df = 1, p value < 0.001 and χ^2 = 343.05, df = 1, p value < 0.001, respectively). Over-triaged patients to ESI level 3 had a significantly lower prevalence of admission and critical outcomes as compared to those appropriately triaged to ESI level 3 (χ^2 = 1184.90 df = 1, p value < 0.001 and χ^2 = 213.04, df = 1, p value < 0.001, respectively). Similarly, patients who were over-triaged to high-acuity ESI levels (1 or 2) had a significantly lower prevalence of admission and critical outcomes as compared to patients appropriately triaged to the same ESI levels (χ^2 = 588.49, df = 1, p value < 0.001 and χ^2 = 126.57, df = 1, p value < 0.001, respectively).
	E. S. S. Velarde et al. [90] (2015; World Congress on Medical Physics and Biomedical Engineering)	General appearance, blood pressure, heart rate, respiratory rate, temperature and Glasgow Coma Scale.	The triage software was able to timely prioritize patients from the output given by the rules of a fuzzy controller.
	D. Azeez et al. [107] (2015; Technology and Health Care)	Chief complaints, patient medical history and vital signs. Vital signs were heart rate, blood pressure, temperature, respiratory rate, oxygen saturation. One of the input features of the model was a free-text space for chief complaint.	Random forest: sensitivity and specificity of 0.98 and 0.89, respectively.
	V. C. Georgopoulos and C. D. Stylios [122] (2015; Simulation and Modeling Methodologies, Technologies and Applications Conference)	Life threatening, limb threatening, patient chief complaint, vital signs, medical history, expected number of resources, patient age, required timely intervention, weakness, additional symptoms, severe pain or distress, patient referred to the ED from outside, behavioral or psychiatric issue, patient medications, hospital or ED discharge inferior to 3 days, patient immune-compromised, alcohol or illicit drug use, no recent change mental state, patient can walk or sit and pre-existing communication/cognitive deficits.	In a case example, the Fuzzy Cognitive Map system assisted the health professional in the prioritization of patients in the ESI levels 3-5.
	E. B. Fields et al. [119] (2013; Expert Systems with Applications)	Vital signs, age, gender, temperature, heart rate, respiration rate, systolic and diastolic blood pressure.	The study recommended the method that estimates utility intervals as the most suitable preference aggregation method for prioritization.
	D. Azeez et al. [33] (2013; SpringerPlus)	Age, gender, airway and breathing, tachychycardia, bradycardia, sever pallor, cold peripheries, tachypnoea, cannot complete in full sentence, one sided limb weakness, slurring of speech, facial asymmetry, sever chest pain, perfuse sweating, altered mental status, sever intractable pain, psychiatric patient irritable, chief complaint, heart rate, respiration rate.	ANN: RMSE, %RMSE and ACC were 0.14, 5.7 and 99%, respectively, for the train set. RMSE, %RMSE and ACC were 0.18, 7.16 and 96.7%, respectively, for the test set. ANFIS: RMSE, %RMSE and ACC were 0.85, 32 and 96%, respectively, for the train set. RMSE, % RMSE and ACC were 1.30, 49.84 and 94%, respectively, for the test set.
	V. C. Georgopoulos and C. D. Stylios [120] (2013; Fuzziness and Medicine: Philosophical Reflections and	All the variables which were used in [122], except for the variable of pre-existing communication/cognitive deficits.	In a case example, the Fuzzy Cognitive Map system assisted the health professional in the prioritization of elderly patients.

(*continued on next page*)

Table 4 (*continued*)

Objectives	Author (Year; Source)	Independent Variables	Results by outcome (Algorithm: performance)
	Application Systems in Health Care - Book chapter)		
	W-T. Lin et al. [66] (2013; Proceedings of World Academy of Science, Engineering and Technology)	Chief complaint, medical history, general appearances, vital signs, symptoms and signs and the results of a physical assessment.	ANN ACC: 95.1%
	S.-T. Wang [70] (2013; Journal of Medical Systems)	Chief complaints, medical history, general appearance, vital signs, symptoms and signs, and physical assessment results. The vital signs consisted of 6 parameters, including breathing, temperature, heart rate, diastolic pressure, systolic pressure and SpO2.	For the anomaly detection, the ACC was: PCA with SVM: 100 %; SVM: 89.2 %; BPNN: 96.71 %. To predict triage priority, the MAPE was: SVR: 3.78 %; BPNN: 5.99 %.
	D. Zmiri et al. [79] (2012; Journal of Evaluation in Clinical Practice)	Age, gender, vital signs (e.g. temperature, heart rate, blood pressure, respiration rate, oxygen saturation and glucose), chief complaints, previous diagnoses known at the time of arrival to the ED, and a letter of referral to the ED.	DT and NB: mean ACC 52.94 ± 5.89%, which was significant (*p*-value < 0.05) when compared to the mean ACC of a random classifier (34.60 ± 2.40%). Allowing for classification deviations of one severity grade led to mean ACC of 85.42 ± 1.42%.
	D. Aziz et al. [118] (2012; International Conference on Intelligent and Advanced Systems (ICIAS))	Primary triage attributes from OPTS: airway and breathing, tachycardia, bradycardia, severe pallor, cold peripheries, tachypnoea, cannot complete full sentences, one sided limb weakness, slurring of speech, facial asymmetry, severe chest pain, altered mental status, polytrauma, severe intractable pain, psychiatric patient irritable and profuse sweating.	ANFIS: ACC 98.4%. The fuzzy C-means method produced fewer rules and needed less processing time to reach the RMSE of 0.127 compared to the fuzzy subtractive clustering method.
	V. Georgopoulos and C. Stylios [124] (2012; IFAC Proceedings Volumes)	Patient chief complaints, vital signs, medical history, expected number of resources, age, required timely intervention, additional symptoms other than chief complaint, severe pain or distress, patient referred to ED from outside, behavioral or psychiatric issue, no additional symptoms to chief complaint, absence of medical history, patient medications, hospital or ED discharge less than 3 days, patient immune-compromised, alcohol or illicit drug use.	Results in [120,122].
	W.-T. Lin et al. [121] (2011; Expert Systems with Applications)	Age, gender, patient type (first visit and revisit), insurance status, period (AM, PM and night), hospital admission, arrival mode, subject (internal, surgery, obstetrics and gynecology, pediatrics, dentistry and psychosomatic medicine), overstay length and expenses.	By two-stage cluster analysis (Ward's method and K-means), it was concluded that the cohort of elderly patients with longer ED LOS, higher consumption of medical expenses were in the group with higher risk of consumption of resources, and the majority were patients of most urgent triage priorities 1 and 2.
	W.-T. Lin et al. [103] (2010; Expert Systems with Applications)	Vital signs: respiration rate, diastolic pressure, systolic pressure, SaO2, heart rate and temperature. Decision parameters: nursing personnel triage, physicians' triage and decision time interval.	By two-stage cluster analysis (Ward's method and K-means) and DT, it was found that heart rate and temperature were important factors to detect abnormal patients. It was also found that abnormal diagnosis was most likely in diseases of pneumonia and cirrhosis.
Predict ED LOS	M. Street et al. [65] (2018; European Journal of Emergency Medicine)	Age, usual accommodation, triage category, arrival by ambulance, arrival overnight, imaging, laboratory investigations, overcrowding, time to be seen by doctor, ED visits with admission and access block relating to ED LOS more than 4 h.	LR AUC in the validation set: 0.80, (Hosmer-Lemeshow p-value 0.36 and prediction MSE 0.18).
	A. Azari et al. [31] (2015; IEEE International Conference on Bioinformatics and Biomedicine (BIBM))	Patient's acuity level, the chief complaint, temperature, respiration rate, heart rate, oxygenation rate, mean arterial blood pressure rate, the department in the ED, and if a speciality team is called.	The sensitivity was very high for the majority class of non admitted patients (96%). However, the increased sensitivity for the majority class was at the cost of much lower sensitivity for the minority class of admitted patients (26%).

(*continued on next page*)

Table 4 (*continued*)

Objectives	Author (Year; Source)	Independent Variables	Results by outcome (Algorithm: performance)
	E. Casalino et al. [112] (2014; Emergency Medicine Journal)	Age, gender, triage acuity level, and final disposition decisions from ED and Observation Unit (OU).	LR for ED - LOS AUC: 0.791 (95% CI 0.788 to 0.794); LR for ED-OU LOS AUC: 0.812 (95% CI 0.81 to 0.814);
	R. Ding et al. [115] (2009; International Conference on Management and Service Science)	Age, gender, mode of arrival, arrival day and time, insurance status, acuity level and chief complaint.	Patients at the 90th percentile waited 7 times longer (98 minutes), took 2.5 times longer to be treated (487 minutes) and boarded 7 times longer (1,122 minutes) compared to patients at the median.

References

[1] Goldhill D, Worthington L, Mulcahy A, Tarling M, Sumner A. The patient-at-risk team: identifying and managing seriously ill ward patients. ANAESTHESIA-LONDON- 1999;54:853–60.
[2] McQuillan P, Pilkington S, Allan A, Taylor B, Short A, Morgan G, Nielsen M, Barrett D, Smith G. Confidential inquiry into quality of care before admission to intensive care. BMJ 1998;316(7148):1853–8.
[3] Carter EJ, Pouch SM, Larson EL. The relationship between emergency department crowding and patient outcomes: a systematic review. Journal of Nursing Scholarship 2014;46(2):106–15.
[4] Guttmann A, Schull MJ, Vermeulen MJ, Stukel TA. Association between waiting times and short term mortality and hospital admission after departure from emergency department: population based cohort study from Ontario, Canada. BMJ 2011;342:d2983.
[5] Guedes HM, Martins JCA, Chianca TCM. Predictive value of the Manchester Triage System: evaluation of patients' clinical outcomes. Revista Brasileira de Enfermagem 2015;68(1):45–51.
[6] Guedes HM, Souza KM, Lima PdO, Martins JCA, Chianca TCM. Relationship between complaints presented by emergency patients and the final outcome. Revista Latino-Americana de Enfermagem 2015;23(4):587–94.
[7] Caroline Gonçales P, Pinto Júnior D, de Oliveira Salgado P, Machado Chianca TC. Relationship between risk stratification, mortality and length of stay in a Emergency Hospital. Investigación y Educación en Enfermería 2015;33(3):424–31.
[8] Gupta JN, Forgionne GA, Mora M. Intelligent decision-making support systems: foundations, applications and challenges. Springer Science & Business Media; 2007.
[9] Merkert J, Mueller M, Hubl M. A Survey of the Application of Machine Learning in Decision Support Systems. ECIS 2015.
[10] Proudlove N, Vadera S, Kobbacy K. Intelligent management systems in operations: a review. Journal of the Operational Research Society 1998;49(7):682–99.
[11] Gottinger HW, Weimann P. Intelligent decision support systems. Decision Support Systems 1992;8(4):317–32.
[12] Elam JJ, Konsynski B. Using artificial intelligence techniques to enhance the capabilities of model management systems. Decision Sciences 1987;18(3):487–502.
[13] Garg AX, Adhikari NK, McDonald H, Rosas-Arellano MP, Devereaux P, Beyene J, Sam J, Haynes RB. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. Jama 2005;293(10):1223–38.
[14] Pombo N, Araújo P, Viana J. Knowledge discovery in clinical decision support systems for pain management: A systematic review. Artificial intelligence in medicine 2014;60(1):1–11.
[15] Miller RA. Why the standard view is standard: people, not machines, understand patients' problems. Journal of Medicine and Philosophy 1990;15(6):581–91.
[16] Shortliffe EH, Cimino JJ. Biomedical informatics: computer applications in health care and biomedicine. Springer Science & Business Media; 2013.
[17] Stone CP. A glimpse at EHR implementation around the world: the lessons the US can learn. The Health Institute for E-Health Policy; 2014.
[18] Linder JA, Ma J, Bates DW, Middleton B, Stafford RS. Electronic health record use and the quality of ambulatory care in the United States. Archives of Internal Medicine 2007;167(13):1400–5.
[19] Nebeker JR, Hoffman JM, Weir CR, Bennett CL, Hurdle JF. High rates of adverse drug events in a highly computerized hospital. Archives of Internal Medicine 2005;165(10):1111–6.
[20] National Research Council, et al. Computational technology for effective health care: immediate steps and strategic directions. National Academies Press; 2009.
[21] Walker J, Pan E, Johnston D, Milsten J, Bates D, Middleton B. The Value of Computerized Provider Order Entry in Ambulatory Settings. Boston (MA): Center for Information Technology Leadership; 2004.
[22] Berlyand Y, Raja AS, Dorner SC, Prabhakar AM, Sonis JD, Gottumukkala RV, Succi MD, Yun BJ. How artificial intelligence could transform emergency department operations. The American journal of emergency medicine 2018;36(8):1515–7.
[23] Kitchenham B, Brereton OP, Budgen D, Turner M, Bailey J, Linkman S. Systematic literature reviews in software engineering - a systematic literature review. Information and Software Technology 2009;51(1):7–15.
[24] Kitchenham BA, Dyba T, Jorgensen M. Evidence-based software engineering. Proceedings of the 26th international conference on software engineering 2004:273–81.
[25] Pereira CR, Pereira DR, Weber SA, Hook C, de Albuquerque VHC, Papa JP. A survey on computer-assisted parkinson's disease diagnosis. Artificial intelligence in medicine; 2018.
[26] Haddaway NR, Collins AM, Coughlin D, Kirk S. The role of google scholar in evidence reviews and its applicability to grey literature searching. PloS one 2015;10(9):e0138237.
[27] Lidal IB, Holte HH, Vist GE. Triage systems for pre-hospital emergency medical services-a systematic review. Scandinavian journal of trauma, resuscitation and emergency medicine 2013;21(1):28.
[28] Belard A, Buchman T, Forsberg J, Potter BK, Dente CJ, Kirk A, Elster E. Precision diagnosis: a view of the clinical decision support systems (CDSS) landscape through the lens of critical care. Journal of clinical monitoring and computing 2017;31(2):261–71.
[29] Teubner DJ, Considine J, Hakendorf P, Kim S, Bersten AD. Model to predict inpatient mortality from information gathered at presentation to an emergency department: The Triage Information Mortality Model (TIMM). Emergency Medicine Australasia 2015;27(4):300–6.
[30] Zlotnik A, Alfaro MC, Pérez MCP, Gallardo-Antolín A, Martínez JMM. Building a Decision Support System for Inpatient Admission Prediction With the Manchester Triage System and Administrative Check-in Variables. CIN: Computers, Informatics, Nursing 2016;34(5):224–30.
[31] Azari A, Janeja VP, Levin S. Imbalanced learning to predict long stay emergency department patients. Bioinformatics and Biomedicine (BIBM), 2015 IEEE International Conference on 2015:807–14.
[32] LaMantia MA, Stewart PW, Platts-Mills TF, Biese KJ, Forbach C, Zamora E, McCall BK, Shofer FS, Cairns CB, Busby-Whitehead J, et al. Predictive value of initial triage vital signs for critically ill older adults. Western Journal of Emergency Medicine 2013;14(5):453.
[33] Azeez D, Ali MAM, Gan KB, Saiboon I. Comparison of adaptive neuro-fuzzy inference system and artificial neutral networks model to categorize patients in the emergency department. SpringerPlus 2013;2(1):416.
[34] Zhang X, Kim J, Patzer RE, Pitts SR, Patzer A, Schrager JD. Prediction of emergency department hospital admission based on natural language processing and neural networks. Methods of information in medicine 2017;56(05):377–89.
[35] Jernite Y, Halpern Y, Horng S, Sontag D. Predicting chief complaints at triage time in the emergency department. NIPS 2013 Workshop on Machine Learning for Clinical Data Analysis and Healthcare 2013.
[36] Thompson DA, Eitel D, Fernandes CM, Pines JM, Amsterdam J, Davidson SJ. Coded chief complaints-automated analysis of free-text complaints. Academic emergency medicine 2006;13(7):774–82.
[37] Handly N, Thompson DA, Li J, Chuirazzi DM, Venkat A. Evaluation of a hospital admission prediction model adding coded chief complaint data using neural network methodology. European Journal of Emergency Medicine 2015;22(2):87–91.
[38] Cramer JS. The origins of logistic regression. 2002.
[39] Dreiseitl S, Ohno-Machado L. Logistic regression and artificial neural network classification models: a methodology review. Journal of biomedical informatics 2002;35(5-6):352–9.
[40] Ruger JP, Lewis LM, Richter CJ. Identifying high-risk patients for triage and resource allocation in the ED. The American Journal of Emergency Medicine 2007;25(7):794–8.
[41] Cameron A, Rodgers K, Ireland A, Jamdar R, McKay GA. A simple tool to predict admission at the time of triage. Emerg Med J 2015;32(3):174–9.
[42] Sun Y, Heng BH, Tay SY, Seow E. Predicting hospital admissions at emergency department triage using routine administrative data. Academic Emergency Medicine 2011;18(8):844–50.
[43] Barfod C, Lauritzen MMP, Danker JK, S“olétormos G, Forberg JL, Berlac PA, Lippert F, Lundstrøm LH, Antonsen K, Lange KHW. Abnormal vital signs are strong predictors for intensive care unit admission and in-hospital mortality in adults triaged in the emergency department-a prospective cohort study. Scandinavian journal of trauma, resuscitation and emergency medicine 2012;20(1):28.
[44] Coslovsky M, Takala J, Exadaktylos AK, Martinolli L, Merz TM. A clinical prediction model to identify patients at high risk of death in the emergency

department. Intensive care medicine 2015;41(6):1029–36.
[45] Graham B, Bond R, Quinn M, Mulvenna M. Using data mining to predict hospital admissions from the emergency department. IEEE Access 2018;6:10458–69.
[46] Ramírez Nafarrate A. Predictive analytics for hospital admissions from the emergency department using triage information. OPENAIRE. 2019.
[47] Hong WS, Haimovich AD, Taylor RA. Predicting hospital admission at emergency department triage using machine learning. PloS one 2018;13(7):e0201016.
[48] Parker CA, Liu N, Wu SX, Shen Y, Lam SSW, Ong MEH. Predicting hospital admission at the emergency department triage: A novel prediction model. The American journal of emergency medicine 2018.
[49] LaMantia MA, Platts-Mills TF, Biese K, Khandelwal C, Forbach C, Cairns CB, Busby-Whitehead J, Kizer JS. Predicting hospital admission and returns to the emergency department for elderly patients. Academic emergency medicine 2010;17(3):252–9.
[50] Plunkett PK, Byrne DG, Breslin T, Bennett K, Silke B. Increasing wait times predict increasing mortality for emergency medical admissions. European Journal of Emergency Medicine 2011;18(4):192–6.
[51] Pereira L, Choquet C, Perozziello A, Wargon M, Juillien G, Colosi L, Hellmann R, Ranaivoson M, Casalino E. Unscheduled-return-visits after an emergency department (ED) attendance and clinical link between both visits in patients aged 75 years and over: a prospective observational study. PloS one 2015;10(4):e0123803.
[52] Hinson JS, Martinez DA, Schmitz PS, Toerper M, Radu D, Scheulen J, de Ramirez SAS, Levin S. Accuracy of emergency department triage using the emergency severity index and independent predictors of under-triage and over-triage in Brazil: a retrospective cohort analysis. International journal of emergency medicine 2018;11(1):3.
[53] Barnes S, Saria S, Levin S. An evolutionary computation approach for optimizing multilevel data to predict patient outcomes. Journal of healthcare engineering 2018;2018.
[54] Ng C-J, Liao P-J, Chang Y-C, Kuan J-T, Chen J-C, Hsu K-H. Predictive factors for hospitalization of nonurgent patients in the emergency department. Medicine 2016;95(26).
[55] Srivilaithon W, Amnuaypattanapon K, Limjindaporn C, Imsuwan I, Daorattanachai K, Dasanadeba I, Siripakarn Y. Predictors of in-hospital cardiac arrest within 24 h after emergency department triage: A case-control study in urban Thailand. Emergency Medicine Australasia; 2019.
[56] Lee DS, Stitt A, Austin PC, Stukel TA, Schull MJ, Chong A, Newton GE, Lee JS, Tu JV. Prediction of heart failure mortality in emergent care: a cohort study. Annals of internal medicine 2012;156(11):767–75.
[57] Jenny MA, Hertwig R, Ackermann S, Messmer AS, Karakoumis J, Nickel CH, Bingisser R. Are mortality and acute morbidity in patients presenting with nonspecific complaints predictable using routine variables? Academic Emergency Medicine 2015;22(10):1155–63.
[58] Ljunggren M, Castrén M, Nordberg M, Kurland L. The association between vital signs and mortality in a retrospective cohort study of an unselected emergency department population. Scandinavian journal of trauma, resuscitation and emergency medicine 2016;24(1):21.
[59] Kim SW, Li JY, Hakendorf P, Teubner DJ, Ben-Tovim DI, Thompson CH. Predicting admission of patients by their presentation to the emergency department. Emergency Medicine Australasia 2014;26(4):361–7.
[60] Henning DJ, Oedorf K, Day DE, Redfield CS, Huguenel CJ, Roberts JC, Sanchez LD, Wolfe RE, Shapiro NI. Derivation and validation of predictive factors for clinical deterioration after admission in emergency department patients presenting with abnormal vital signs without shock. Western Journal of Emergency Medicine 2015;16(7):1059.
[61] Riordan JP, Dell WL, Patrie JT. Can patient variables measured on arrival to the emergency department predict disposition in medium-acuity patients? The Journal of emergency medicine 2017;52(5):769–79.
[62] Raita Y, Goto T, Faridi MK, Brown DF, Camargo CA, Hasegawa K. Emergency department triage prediction of clinical outcomes using machine learning models. Critical Care 2019;23(1):64.
[63] Hong W, Earnest A, Sultana P, Koh Z, Shahidah N, Ong MEH. How accurate are vital signs in predicting clinical outcomes in critically ill emergency department patients. European Journal of Emergency Medicine 2013;20(1):27–32.
[64] Dj”arv T, Castrén M, Mårtenson L, Kurland L. Decreased general condition in the emergency department: high in-hospital mortality and a broad range of discharge diagnoses. European Journal of Emergency Medicine 2015;22(4):241–6.
[65] Street M, Mohebbi M, Berry D, Cross A, Considine J. Influences on emergency department length of stay for older people. European Journal of Emergency Medicine 2018;25(4):242–9.
[66] Lin W-T, Jou Y-T, Wu Y-C, Hsiao Y-D. Data mining applied to the predictive model of triage system in emergency department. Proceedings of World Academy of Science, Engineering and Technology 2013:1789. no. 78, Citeseer.
[67] Ro YS, Shin SD, Song KJ, Cha WC, Cho JS. Triage-based resource allocation and clinical treatment protocol on outcome and length of stay in the emergency department. Emergency Medicine Australasia 2015;27(4):328–35.
[68] Vapnik V. The nature of statistical learning theory. Springer science & business media; 2013.
[69] Cortes C, Vapnik V. Support-vector networks. Machine learning 1995;20(3):273–97.
[70] Wang S-T. Construct an optimal triage prediction model: A case study of the emergency department of a teaching hospital in Taiwan. Journal of Medical Systems 2013;37(5):9968.
[71] Olivia D, Nayak A, Balachandra M. Machine learning based electronic triage for emergency department. International Conference on Applications and Techniques in Information Security 2018:215–21.
[72] Ong MEH, Ng CHL, Goh K, Liu N, Koh ZX, Shahidah N, Zhang TT, Fook-Chong S, Lin Z. Prediction of cardiac arrest in critically ill patients presenting to the emergency department using a machine learning score incorporating heart rate variability compared with the modified early warning score. Critical Care 2012;16(3):R108.
[73] Cristianini N, Shawe-Taylor J, et al. An introduction to support vector machines and other kernel-based learning methods. Cambridge university press; 2000.
[74] Sch“olkopf B, Smola AJ, Bach F, et al. Learning with kernels: support vector machines, regularization, optimization, and beyond. MIT press; 2002.
[75] Jinglin Y, Li H-X, Yong H. A probabilistic SVM based decision system for pain diagnosis. Expert Systems with Applications 2011;38(8):9346–51.
[76] Tan KC, Teoh EJ, Yu Q, Goh K. A hybrid evolutionary algorithm for attribute selection in data mining. Expert Systems with Applications 2009;36(4):8616–30.
[77] Kotsiantis SB, Zaharakis ID, Pintelas PE. Machine learning: a review of classification and combining techniques. Artificial Intelligence Review 2006;26(3):159–90.
[78] Maron ME. Automatic indexing: an experimental inquiry. Journal of the ACM (JACM) 1961;8(3):404–17.
[79] Zmiri D, Shahar Y, Taieb-Maimon M. Classification of patients by severity grades during triage in the emergency department using data mining methods. Journal of Evaluation in Clinical Practice 2012;18(2):378–88.
[80] Kuncheva LI. Fuzzy classifiers. Scholarpedia 2008;3(1):2925.
[81] Zadeh LA. Fuzzy sets. Information and control 1965;8(3):338–53.
[82] Fathi-Torbaghan M, Meyer D. Medusa: a fuzzy expert system for medical diagnosis of acute abdominal pain. Methods of information in medicine 1994;33(05):522–9.
[83] Binaghi E, Gallo I, Ghiselli C, Levrini L, Biondi K. An integrated fuzzy logic and web-based framework for active protocol support. International Journal of Medical Informatics 2008;77(4):256–71.
[84] Zwick R, Wallsten TS. Combining stochastic uncertainty and linguistic inexactness: theory and experimental evaluation of four fuzzy probability models. International Journal of Man-Machine Studies 1989;30(1):69–111.
[85] Yaguinuma CA, Santos MT, Camargo HA, Nogueira TM. A meta-ontology approach for representing vague linguistic terms and fuzzy rules for classification in ontologies. 2010 14th IEEE International Enterprise Distributed Object Computing Conference Workshops 2010:263–71.
[86] Dombi J. Membership function as an evaluation. Fuzzy sets and systems 1990;35(1):1–21.
[87] Abbasi M, Kashiyarndi S. Clinical decision support systems: A discussion on different methodologies used in health care. Marlaedalen University Sweden; 2006.
[88] Nilashi M, Ibrahim O, Ahani A. Accuracy improvement for predicting parkinson's disease progression. Scientific reports 2016;6:34181.
[89] Cherifi H, Zain JM, El-Qawasmeh E. Digital Information and Communication Technology and Its Applications: International Conference, DICTAP 2011, Dijon, France, June 21-23, 2011. Proceedings, Vol. 166 2011.
[90] Velarde ESS, Sotelo-de Ávila AA, Rico-Asención IO, Alemán NR, González R-S, Ramírez-Sotelo MG, Cabrera-Llanos AI. Fuzzy-state machine for Triage priority classifier in emergency room. World Congress on Medical Physics and Biomedical Engineering, June 7-12, 2015. Springer; 2015. p. 1488–91.
[91] Soufi MD, Samad-Soltani T, Vahdati SS, Rezaei-Hachesu P. Decision support system for triage management: A hybrid approach using rule-based reasoning and fuzzy logic. International journal of medical informatics 2018;114:35–44.
[92] Andriulli A, Grossi E, Buscema M, Festa V, Intraligi N, Dominici P, Cerutti R, Perri F, Group NLS, et al. Contribution of artificial neural networks to the classification and treatment of patients with uninvestigated dyspepsia. Digestive and liver disease 2003;35(4):222–31.
[93] Das A, Ben-Menachem T, Cooper GS, Chak A, Sivak Jr. MV, Gonet JA, Wong RC. Prediction of outcome in acute lower-gastrointestinal haemorrhage based on an artificial neural network: internal and external validation of a predictive model. The Lancet 2003;362(9392):1261–6.
[94] Hochreiter S, Schmidhuber J. Long short-term memory. Neural computation 1997;9(8):1735–80.
[95] Gligorijevic D, Stojanovic J, Satz W, Stojkovic I, Schreyer K, Del Portal D, Obradovic Z. Deep attention model for triage of emergency department patients. Proceedings of the 2018 SIAM International Conference on Data Mining 2018:297–305.
[96] Tsai F-S, Weng Y-M, Ng C-J, Lee C-C. Embedding stacked bottleneck vocal features in a LSTM architecture for automatic pain level classification during emergency triage. 2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII) 2017:313–8.
[97] Lorena AC, Jacintho LF, Siqueira MF, De Giovanni R, Lohmann LG, De Carvalho AC, Yamamoto M. Comparing machine learning classifiers in potential distribution modelling. Expert Systems with Applications 2011;38(5):5268–75.
[98] Meyfroidt G, G”uiza F, Ramon J, Bruynooghe M. Machine learning techniques to examine large patient databases. Best Practice & Research Clinical Anaesthesiology 2009;23(1):127–43.
[99] Kennedy R, Harrison R, Burton A, Fraser H, Hamer W, MacArthur D, McAllum R, Steedman D. An artificial neural network system for diagnosis of acute myocardial infarction (AMI) in the accident and emergency department: evaluation and comparison with serum myoglobin measurements. Computer Methods and Programs in Biomedicine 1997;52(2):93–103.
[100] Kononenko I. Machine learning for medical diagnosis: history, state of the art and perspective. Artificial Intelligence in medicine 2001;23(1):89–109.
[101] Jang J-S. Self-learning fuzzy controllers based on temporal backpropagation. IEEE Transactions on neural networks 1992;3(5):714–23.
[102] Bodyanskiy Y, Dolotov A. Hybrid systems of computational intelligence evolved from self-learning spiking neural network. Methods and Instruments of Artificial Intelligence 2010;17.

[103] Lin W-T, Wang S-T, Chiang T-C, Shi Y-X, Chen W-Y, Chen H-M. Abnormal diagnosis of Emergency Department triage explored with data mining technology: An Emergency Department at a Medical center in Taiwan taken as an example. Expert Systems with Applications 2010;37(4):2733–41.
[104] Yang C-C, Lin W, Chen H, Shi Y. Improving scheduling of emergency physicians using data mining analysis. Expert Systems with Applications 2009;36(2):3378–87.
[105] Li H, Wang M, Zhou X, Zhao J. An interval set model for learning rules from incomplete information table. International Journal of Approximate Reasoning 2012;53(1):24–37.
[106] Breiman L. Random forests. Machine learning 2001;45(1):5–32.
[107] Azeez D, Gan K, Ali M, Ismail M. Secondary triage classification using an ensemble random forest technique. Technology and Health Care 2015;23(4):419–28.
[108] Levin S, Toerper M, Hamrock E, Hinson JS, Barnes S, Gardner H, Dugas A, Linton B, Kirsch T, Kelen G. Machine-learning-based electronic triage more accurately differentiates patients with respect to clinical outcomes compared with the emergency severity index. Annals of Emergency Medicine 2017.
[109] Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: data mining, inference, and prediction. springer series in statistics. 2009.
[110] Friedman JH. Greedy function approximation: a gradient boosting machine. Annals of statistics 2001:1189–232.
[111] Chen T, Guestrin C. Xgboost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining 2016:785–94.
[112] Casalino E, Wargon M, Peroziello A, Choquet C, Leroy C, Beaune S, Pereira L, Bernard J, Buzzi J-C. Predictive factors for longer length of stay in an emergency department: a prospective multicentre study evaluating the impact of age, patient's clinical acuity and complexity, and care pathways. Emerg Med J 2014;31(5):361–8.
[113] Handly N, Thompson D, Venkat A. Derivation and validation of a hospital admission prediction model adding coded chief complaint to demographic, emergency department operational and patient acuity data available at emergency department triage using neural net methodology. Annals of Emergency Medicine 2013;62(4):S138.
[114] Hendin A, Eagles D, Myers V, Stiell IG. Characteristics and outcomes of older emergency department patients assigned a low acuity triage score. Canadian Journal of Emergency Medicine 2018;20(5):762–9.
[115] Ding R, McCarthy ML, Lee J, Desmond JS, Zeger SL, Aronsky D. Predicting emergency department length of stay using quantile regression. 2009 International Conference on Management and Service Science 2009:1–4.
[116] Ang E, Kwasnick S, Bayati M, Plambeck EL, Aratow M. Accurate emergency department wait time prediction. Manufacturing & Service Operations Management 2015;18(1):141–56.
[117] Sun Y, Teow KL, Heng BH, Ooi CK, Tay SY. Real-time prediction of waiting time in the emergency department, using quantile regression. Annals of emergency medicine 2012;60(3):299–308.
[118] Aziz D, Ali MM, Gan K, Saiboon I. Initialization of adaptive neuro-fuzzy inference system using fuzzy clustering in predicting primary triage category. Intelligent and Advanced Systems (ICIAS), 2012 4th International Conference on, Vol. 1 2012:170–4.
[119] Fields EB, Okudan GE, Ashour OM. Rank aggregation methods comparison: A case for triage prioritization. Expert Systems with Applications 2013;40(4):1305–11.
[120] Georgopoulos VC, Stylios CD. Fuzzy cognitive map decision support system for successful triage to reduce unnecessary emergency room admissions for the elderly. Fuzziness and Medicine: Philosophical Reflections and Application Systems in Health Care 2013:415–36.
[121] Lin W, Wu Y, Zheng J, Chen M. Analysis by data mining in the emergency medicine triage database at a Taiwanese regional hospital. Expert Systems with Applications 2011;38(9):11078–84.
[122] Georgopoulos VC, Stylios CD. Supervisory Fuzzy Cognitive Map Structure for Triage Assessment and Decision Support in the Emergency Department. Simulation and Modeling Methodologies, Technologies and Applications 2015:255–69.
[123] Georgopoulos VC, Stylios CD. Fuzzy cognitive maps for decision making in triage of non-critical elderly patients. 2017 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS) 2017:225–8.
[124] Georgopoulos VC, Stylios CD. Introducing fuzzy cognitive maps for developing decision support system for triage at emergency room admissions for the elderly. IFAC Proceedings Volumes 2012;45(18):484–9.
[125] Wang Y-M, Parkan C. A minimax disparity approach for obtaining OWA operator weights. Information Sciences 2005;175(1):20–9.
[126] Wang Y-M, Luo Y, Hua Z. Aggregating preference rankings using OWA operator weights. Information Sciences 2007;177(16):3356–63.