THINK INDIA JOURNAL

ISSN:0971-1260

Vol-22-Issue-10-November-2019

Predicting Hospital Admission For The Emergency Department Using C4.5 Classifier

Mr. Gajendran

M.Phil Scholar, Department of Computer Science and Applications,

Rathnavel Subramaniam College of Arts and Science.

Mr G Purusothaman

Associate Professor,

Department of Computer Science and Applications, Rathnavel Subramaniam College of Arts and Science.

Dr G Pandiyan

Associate Professor,

Department of Computer Science and Applications, Rathnavel Subramaniam College of Arts and Science.

Abstract: Emergency Department (ED) in hospitals has primary function to treat patients who are suffering wounds which may prompt extreme complications or severe critical illness. Emergency departments are not for treating patients with normal ongoing consideration. For guaranteeing that the sickest patients get seen initial, a sorting system called triage process is utilized to classify patients. Triage is ordinarily performed by an individual from the nursing staff dependent on the patient's demographics, chief objection, and imperative signs. Consequently, the patient is seen by a therapeutic supplier who makes the underlying consideration plan and at last prescribes a mien, which this examination points of confinement to hospital confirmation or discharge. Prediction models in medication look to improve patient mind and increment strategic productivity. Early identification of ED patients who are probably going to require confirmation may empower better advancement of hospital assets through improved comprehension of ED patient blends. It is progressively comprehended that ED crowding is associated with more unfortunate patient results. Notice of executives and inpatient groups with respect to potential confirmations may help lighten this issue. From the point of view of patient consideration in the ED setting, a patient's probability of confirmation may fill in as an intermediary for keenness, which is utilized in a few downstream choices, for example, bed position and the requirement for emergency mediation. For this we propose C4.5 classifier algorithm of information mining systems in distinguishing appropriate answer for anticipate hospital confirmations from the emergency department.

***Keywords**: Data Mining, Medical Data, Emergency Department, Prediction and C4.5 classifier*

I. INTRODUCTION

Emergency department (ED) in hospitals has main function to treat patients who are suffering from injuries which may lead to severe complications or acute serious illness. Emergency departments are not for treating patients with regular ongoing care. EDs represent the largest source of hospital admissions. Upon appearance to the ED, patients are first arranged by sharpness to organize people requiring pressing medical intercession. For guaranteeing that the most broken-down patients get seen initial, an arranging instrument called triage process is utilized to sort patients. Along these lines, the patient is seen by a medical supplier who makes the underlying care plan and at last prescribes a demeanour, which this examination breaking points to hospital admission or discharge. According to the patient’s symptoms, the

ISSN:0971-1260

Vol-22-Issue-10-November-2019

patient is categorized into three levels of triage process. In this, critical patients are given first importance and non-critical patients are given a waiting time according to their symptoms to get treated by doctors. Patients waiting in the EDs lead to crowding and this may have negative impacts on management, patients, etc [1]. Therefore, there is a need to explore methods which will help to improve patient flow, prevent overcrowding and reduce waiting time involved in triage process. Identifying patients which are at high risk of getting admissions from emergency department to hospital will help to reduce the crowding and help in resource management. This can be done with the help of different machine learning techniques by predicting the patient's admission. Previous works predicting ED admission have applied both rule-based methods and machine learning techniques. The goal of this work is to evaluate the performance of a machine learning (ML) to predict hospital admission and patient data available at triage [2].

Machine learning algorithms are used more often within the area of health care. Recent studies have shown that machine learning algorithms can be helpful for disease detection, patient care, and community services to name a few. Since Machine learning models have been used earlier on other parts of health care it can also be used to perform the triaging process at the ED. The appearance of machine learning models has demonstrated guarantee to improve predictive capacity in different conditions (e.g., sepsis, spontaneous exchanges to emergency unit. These methodologies offer preferences in that they represent high-request, non-straight associations among predictors and addition increasingly stable prediction. Ongoing investigations have detailed that the use of machine learning models may give high predictive capacity at ED triage in chose patient populaces and settings. Machine learning approaches have stood out due to their better capacity than predict patient results contrasted and customary methodologies in different settings and malady conditions (e.g., sepsis and spontaneous exchanges to the emergency unit) [4]. The benefits of machine learning approaches incorporate their capacity to process complex nonlinear connections among predictors and yield progressively stable predictions.

Prediction models in medication try to improve patient care and increment calculated productivity. For instance, prediction models for sepsis or intense coronary disorder are intended to caution suppliers of possibly perilous conditions, while models for hospital use or patient-stream empower asset advancement on a frameworks level. Early ID of ED patients who are probably going to require admission may empower better enhancement of hospital assets through improved comprehension of ED patient blends. It is progressively comprehended that ED crowding is connected with more unfortunate patient results. Notice of executives and inpatient groups with respect to potential admissions may help ease this issue [5, 6]. From the point of view of patient care in the ED setting, a patient's probability of admission may fill in as an intermediary for sharpness, which is utilized in a few downstream decisions, for example, bed position and the requirement for crisis mediation. Various earlier investigations have tried to predict hospital admission at the hour of ED triage. Most models just incorporate data gathered at triage, for example, socioeconomics, crucial signs, boss objection, nursing notes, and early diagnostics, while a few models incorporate extra highlights, for example, hospital utilization measurements and past medical history. The objective of this paper is to find the nearest hospital in emergency and recommend the hospital by verifying the availability of the bed status and also know about the nurse status.

II. LITERATURE REVIEW

Goto, Tadahiro, Carlos A. Camargo et al [7] examination of broadly agent data of kids showing to the ED, by utilizing data routinely accessible at the hour of triage, we found that the utilization of machine learning ways to deal with ED triage improved the discriminative capacity to predict clinical and attitude results contrasted and the ordinary triage approach. Also, the machine learning approaches accomplished a high affectability for predicting the basic care result. In particular, these methodologies would lessen the quantity of under triaged basically sick kids in the ordinary triage levels 3 to 5 (i.e., youngsters who might be missed by regular methodologies). Moreover, while customary methodologies may assist clinicians with bettering recognize kids who require hospitalization, the machine learning approaches had a higher explicitness for predicting the hospitalization result, which would maintain a strategic distance from over triaging youngsters who are less sick and may not require broad ED assets. Albeit outer planned approval is required, our discoveries present a chance to apply propelled prediction ways to deal with help the clinician's ED triage decision making, which may, thus, accomplish progressively precise clinical care and ideal asset portion.

Charm Suk Hong et al [8] predict patient aura by utilizing the ED supplier's earlier decision as our actual mark. In doing as such, we can't address the propriety of individual clinical decisions. This and comparative examinations would profit by further investigation into a highest quality level measurement for hospital admission. Studies have recommended that such a measurement will stay subtle. Nonetheless, we expect that future work will adjust reaction factors of enthusiasm with patient-situated results to make an institutionalized measurement for hospital admission. Future investigations may attempt to modify the reaction variable for released patients who came back to the ED the following day to get conceded and for conceded patients who all things considered didn't require admission.

Ellahham, S., and N. Ellahham [9] AI has been applied to predicting the progression of patients into the crisis division; streamlining patient stream to hospital; observing patients in ward and crisis office and predicting the accessibility of bed in-patients. Early recognizable proof of patients in the crisis office requiring admission may maybe help in improving the hospital assets. Different gauging techniques that have been utilized in predicting patient stream incorporate direct relapse, exponential smoothing, time arrangement relapse, and fake neural system. Bayesian system is a probabilistic model.

Patricio Wolff et al [10] presents an investigation configuration itemizing how our examination was directed. This approach might be moved to other multi-class order issues where there is a record of ultimate results after the decision, which permits the assessment of the exhibition of new mixes of new (and progressively objective) marks to prepare machine learning models. Their outcomes show an effective experimentation of various machine learning techniques who have an ongoing enthusiasm for logical research and has a colossal potential to be utilized in genuine clinical settings as a Triage decision bolster device. We have tried a few predictive ML devices, finding that Naive Bayes (NB) is the heartiest methodology regarding meeting per class and reasonableness of death result. Moreover, we found that Random Forest displayed a superior AUC, precision, PPV and explicitness in clinical result predictive limit of high seriousness classes than different models. This outcome outflanks our master-based models and customary Triage models approved as of late in paediatric patients.

Twinkle L. Chauhan et al [11] Overcrowding in Emergency Departments greatly affects the patients alongside negative results. In this manner, it gets imperative to discover creative techniques to improve patient stream and anticipate stuffing. This should be possible with the assistance of various machine learning techniques to predict the hospital admissions from ED. This will help in asset arranging of hospital also. The paper comprises of study of the various strategies utilized for prediction. The paper likewise comprises the execution aftereffect of three unique strategies for the machine learning techniques alongside its outcome's examination.

III. PROBLEM DEFINITION

By analysing the existing work, it has been estimated that 40% of patients appearing at EDs have non-pressing issues. This prompt packed sitting areas and long holding up times. Thus, patients with extreme

ISSN:0971-1260

Vol-22-Issue-10-November-2019

earnest care needs are in danger of not being treated on scheduleand separating by Triage is a multiclass problem, it leads to many tensions and constructing the triage system to predict provide more consistency problem [12]. Admitting the patient in hospital with not availability of room provide great impact to the patients and to the hospitals. Hence a problem in the triaging process is the ED nurse's subjective assessment of patients. This is a problem since it is hard to know the correct assessment of patient when there is a disagreement amongst ED nurses.

IV PROPOSED WORK

Emergency department in hospitals has very important functions such as to treat patients who are suffering from injuries which may lead to severe complications or acute serious illness. Emergency departments are not for treating patients with regular ongoing care. It's for ensuring that the sickest patients get seen first, a sorting mechanism named triage process is used to categorize patients. According to the patient's symptoms, the patient is categorized into three levels of triage process. In this, critical patients are given first importance and non-critical patients are given a waiting time according to their symptoms to get treated by doctors [13]. In many existing work dealing with the application of ML to the construction of Triage systems try to predict the actual Triage decisions by analysing the details with the old dataset hence it is not possible one to predict the hospital availability by verifying with the previous week, day, month dataset. Hence, we want the current or present situation of hospital atmosphere to avoid many emergency situations. We propose a method to predict the hospital environment availability to the emergency patient and use the recommendation technique for providing the suggestion to the patient family by approaching the nearest hospital availability. For this we use C4.5 decision tree classifier algorithm of data mining prediction algorithm.

a. Data Pre-processing

Data pre-processing is one of the first and basic advance to data mining or data investigation. The aftereffects of data pre-processing are legitimately inputted to mining model and got the outcomes. A decent data source can expand the precision of mining, yet in addition raise the productivity of algorithm significantly. All for all situation, data pre-processing alludes to as data cleaning, data incorporate, data change, data decrease, et al., handled before the implementation of data mining algorithm. In this stage data were collected from the hospital in daily manner, the data fields like hospital name, total number of rooms, available rooms, doctor's availability, nurse's availability will be gathered and stored in database. After getting these databases, data pre-processing will be done for which all the missing values, null values, noise data, duplicate value are removed for proceeding with perfect dataset for further process. Here anomalies are removed, and duplicate records are eliminated [14]. Then, numeric attributes are discretized by in a finite number of intervals. After that, a verification method is used to select the hospital attributes. It evaluates the missing values s still available or not for preferring sets of attributes that are highly correlated with the given class. The fact that many attributes depend on one another often disproportionately influences the accuracy of pre-processing.

b. Triage processing

After pre-processing of data then it will go to the stage of triage process. As the patient arrives in emergency department, the process called triage is carried out. If they arrived patient is critical in condition, then two possibilities are there. First, according to the condition we find the nearest hospital availability to do the treatment and in second possibilities is to verify the staff availability in the hospital to do the treatment. If patient can get treated with the medicine or injections then, the patient is treated with that or if the signs and symptoms are not such then patient goes through surgery. This can be done by using the triage process [15]. The model works such that, as soon as the patient arrives in the emergency department, a triage process is carried out of the patient by the casualty officer and in the meantime the history of the patient is being checked. If old patient is there, then according to the medical history of the patient, the officer directs the

ISSN:0971-1260

Vol-22-Issue-10-November-2019

patient to the doctor as the records contain the complete history such as last time when the patient got admitted, what disease does that patient is suffering, etc [16]. As the patient is being treated by the doctor and if it's said that he will be getting admitted to the main hospital then, in that time the inpatient bed is made ready for that patient. If the patient is new then, its records are added to the database of hospital patients and triage is done.

c. Feature Extraction

Information extracted from large data setsof ambulance service containing dispatch and find the needed attributes. Hospital records were extracted from the regional Electronic Medical Records (EMR) system. Here several features are extracted for the model [17]. Such features are selected which are important and which helps in predicting. For providing this mechanism here we use 10 attributes, like date, area name, hospital name, hospital distance, total number of rooms, available rooms, doctor's availability, nurse's availability, doctor status, nurse status were used and here main attributes are selected. Here five main attributes are selected. a) Patient name. b) ED-LEVEL: Hospital services providing immediate initial evaluation and treatment to patients on a 24hrs basis. c) EMSA-TRAUMA-LEVEL: Emergency Medical d) ACUITY: Emergency Department type of visit. e) ADMISSION-TYPE-ED: Total Emergency Department visits like inpatient or outpatient admission.

d. Predicting Hospital Admission

Prediction models in medication look to improve patient care and increment calculated productivity. For instance, prediction models for sepsis or intense coronary disorder are intended to caution suppliers of conceivably perilous conditions, while models for hospital use or patient-stream empower asset improvement on a frameworks level [18]. The predictors for machine learning models were looked over routinely accessible data at ED triage utilizing from the earlier information. Early identification of ED patients who are probably going to require admission may empower better streamlining of hospital assets through improved comprehension of ED patient blends. Presently after the occasions preparing and assessing the model, it is prepared forthe prediction reason where outside data is given as info [19]. The predictors for machine learning models were browsed routinely accessible data at ED triage utilizing from the earlier information. We utilized nursing triage notes, in isolation from other data, to predict ED patient disposition. Triage notes, which may reflect the clinical impression of experienced ED nurses and provide recommending to the patient family. To predicting the process here we apply C4.5 classifier algorithm. C4.5 classification implementation adopting a multiclass strategy is embraced, and the awkwardness of data is likewise considered. The classification algorithm utilized in this paper is C4.5 which is an improvement of the decision tree algorithm ID3 [20]. The algorithm has a similar working standard, yet the figuring of data assortment is in an unexpected way. In ID3, the learning process is done with reference to the calculation of the data collection. The calculation of the data collection in ID3 is same as thecalculation of the data collectionin the feature selection process withthe Collected information (CI) as shown in Equations (1)–(2). In the C4.5 algorithm, the learning process uses the ID3 normalized data collection, is written as

$$\text{Collected Ratio (S, P)} = \frac{\text{Collected Data (S,P)}}{\text{Split data (S,P)}} \text{----- (1)}$$

$$\text{Split Data (S, P)} = \sum_{i=1}^{0} \frac{Sv}{S} \log_2 \left(\frac{Sv}{S}\right) \text{---- (2)}$$

Where S = Sample data collected from hospitals, V= Attribute Value and P= Possible attribute

The method used to assess or validate the accuracy of the model of the proposed system to cross validation. Attribute selection is carried out using the following procedure:

Step 0:CI is calculated for all attributes, based on Equation (2)

Step1: All features are extracted by CIto proceed further process

ISSN:0971-1260

Vol-22-Issue-10-November-2019

Step2: All CIsare noted for N attributes of the training data

$$\text{Total Collected data} = \sum_{i=1}^{N} \text{Collected Data}\,(S, P_i)$$

Step3: The Total data Values (W) are calculated for each attribute P_i, where i = 1,2, …, N

$$W(P_i) = \frac{\text{Collected Data (S,P)}}{\text{Total data}}$$

Step4: Repetitions are performed to add attribute total data,

Wn = 0;

FOR i = 1 to N

Wn = Wn + W($P_{i,}$)

IF Wn ≥ Threshold THEN

go to Step-5

ENDIF

i = i + 1

ENDFOR

Step-5: Attribute number i is selected, which attribute to 1,2 i

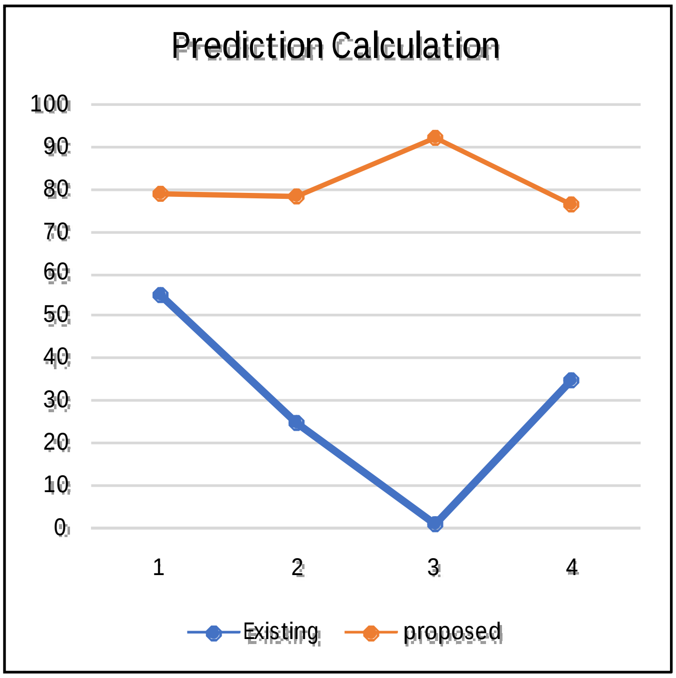
This is used to analyse the variables by grouping and altogether, thus considering not only the contacts among variables within each category but also the relations between diverse classes. This algorithm uses an outfit learning approach for classification while training process by creating number of decision trees. Although the fact that there are different variables that may impact execution of the model, it is conceivable that nursing triage notes have naturally constrained capacity to predict future ED attitude, despite advancement of the predictive model.

V EXPERIMENTAL RESULT

Nowadays, with the growing demand for emergency medical care, the management of hospital emergency departments (EDs) has become increasingly important. The management of patient influx is one of the most crucial problems in Eds. This section shows the experimental results for the emergency departments (ED) to predict and recommend the hospital. To obtain reliable results, the extracted knowledge should be evaluated by a comparison of results obtained with various supervised classification methods and using several measures. In this work, we will apply C4.5, a learning algorithm that will make classification and prediction on a database to provide better suggestion for the patients. Here our implementation work was done by using .NET environment.

Dataset

The dataset used in this study consists of information about 200 patients visits to the ED. The data was gathered between March 2019 to July 2019 in three different EDs located in the Tamilnadu India. For each patient there is 10 different variables that has been collected from the triage evaluation performed by ED predictions.



THINK INDIA JOURNAL

ISSN:0971-1260

Vol-22-Issue-10-November-2019

Chart 1: Comparison of prediction

Predictive modelling is whether additional training samples will improve performance or whether a model has reached its maximum performance in its features. Chart 1 shows the highest prediction mechanism when compared to existing.

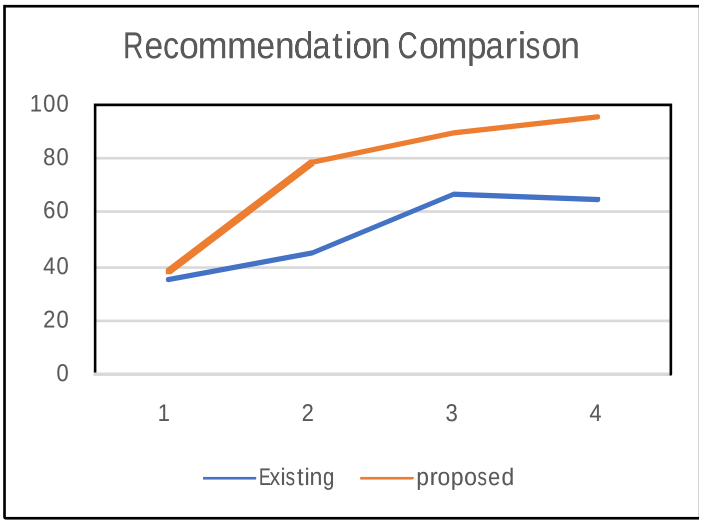


Chart 2: Comparison of Recommendation

Recommendation works very well in this work, the differentiate was shown in chart 2.

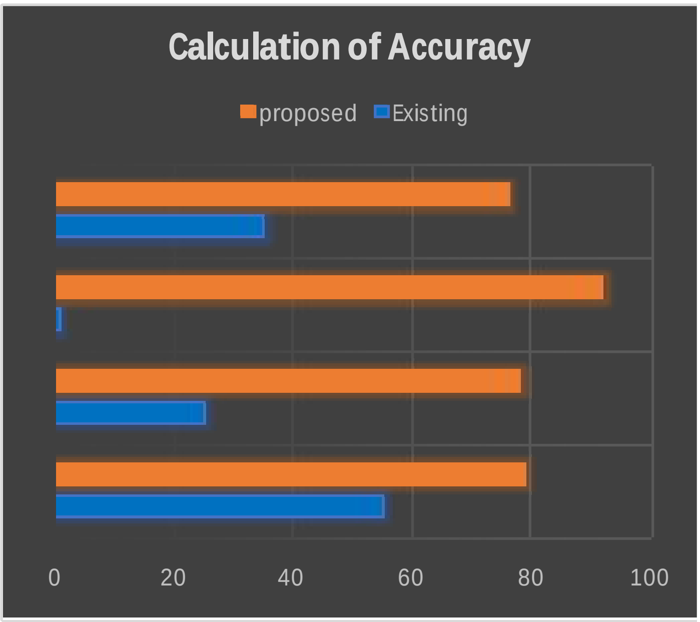


Chart 3: Calculation of Accuracy

Here the accuracy of our proposed work was shown in chart 3. The calculation was given below

Here the accuracy of the data is a rate of correct classification defined by $\frac{TP+TN}{TP+TN+FP+FN}$

Exactness is a rate of true positive classification defined by $\frac{TP}{TP+FP}$

Recall is the evaluation and ranking of each sample based on positive class defined by $\frac{TP}{TP+FN}$

Statistic: measures the agreement of prediction with the true class labels. A score value of 1.0 signifies complete agreement, and a value greater than 0 means that the classifier is doing better than pure random behaviour.

VI. CONCLUSION

ED plays an important one in medical field,Early identification of patients who will be admitted to the ED could help manage workflow through requesting hospital beds earlier during the visit.Development of emergency department (ED) triage frameworks that precisely separate and organize basically sick from stable patients stays testing. We utilized machine learning models to predict the hospital for the ED patient. This paper uses the data mining algorithm of machine learning technique to easily identify the hospital availability and the status of the doctor and nurse's presence state. Our system provides many advantages to the ED patients by approaching as a priori knowledge. For this we use C4.5 classification algorithm to provide prediction and recommendation model.

VII. REFERENCES

[1]. Raita, Yoshihiko et al. "Emergency department triage prediction of clinical outcomes using machine learning models". In: Critical Care 23.1 (2019), p. 64.

[2]. McHugh, Megan et al. "More patients are triaged using the Emergency Severity Index than any other triage acuity system in the United States". In: Academic Emergency Medicine 19.1 (2012), pp. 106–109.

ISSN:0971-1260
Vol-22-Issue-10-November-2019

[3]. Manogaran, Gunasekaran and Lopez, Daphne. “A survey of big data architectures and machine learning algorithms in healthcare”. In: International Journal of Biomedical Engineering and Technology 25.2-4 (2017), pp. 182–211.

[4]. Levin, Scott et al. “Machine-learning-based electronic triage more accurately differentiates patients with respect to clinical outcomes compared with the emergency severity index”. In: Annals of emergency medicine 71.5 (2018), pp. 565– 574.

[5]. Jernite, Yacine, Yoni Halpern, Steven Horng, and David Sontag. "Predicting chief complaints at triage time in the emergency department." In NIPS 2013 Workshop on Machine Learning for Clinical Data Analysis and Healthcare. 2013.

[6]. Sterling, Nicholas W., Rachel E. Patzer, Mengyu Di, and Justin D. Schrager. "Prediction of Emergency Department Patient Disposition Based on Natural Language Processing of Triage Notes." International Journal of Medical Informatics (2019).

[7]. Goto, Tadahiro, Carlos A. Camargo, Mohammad Kamal Faridi, Robert J. Freishtat, and Kohei Hasegawa. "Machine learning–based prediction of clinical outcomes for children during emergency department triage." JAMA network open 2, no. 1 (2019): e186937-e186937.

[8]. Hong, Woo Suk, Adrian Daniel Haimovich, and R. Andrew Taylor. "Predicting hospital admission at emergency department triage using machine learning." PloS one 13, no. 7 (2018): e0201016.

[9]. Ellahham, S., and N. Ellahham. "Use of Artificial Intelligence for Improving Patient Flow and Healthcare Delivery." J Comput Sci Syst Biol 12, no. 303 (2019): 2.

[10]. Wolff, Patricio, Sebastián A. Ríos, and Manuel Graña. "Setting up standards: A methodological proposal for pediatric Triage machine learning model construction based on clinical outcomes." Expert Systems with Applications 138 (2019): 112788.

[11]. Twinkle L. Chauhan et al, “Hospital Admission Prediction: A Technology Survey”, IRJET, Volume: 06 Issue: 05 | May 2019

[12]. Graham, Byron, Raymond Bond, Michael Quinn, and Maurice Mulvenna. "Using data mining to predict hospital admissions from the emergency department." IEEE Access 6 (2018): 10458-10469.

[13]. Chauhan Twinkle, Takalikar Mukta, Gupta Anilkumar, “Hospital Admission Prediction: A Technology Survey”, International Research Journal of Engineering and Technology, vol. 6, pp. 5078-5081, 2019.

[14]. Sun, Yan and Heng, Bee Hoon and Tay, SeowYian and Seow, Eillyne,”Predicting hospital admissions at emergency department triage using routine administrative data,” Academic emergency Medicine., vol. 18, pp. 844–850, 2011.

[15]. Spangler, Douglas Nils, Thomas Hermansson, David Smekal, and Hans Blomberg. "A validation of machine learning-based risk scores in the prehospital setting." medRxiv (2019): 19007021.

[16]. Thitiprayoonwongse, D., Suriyaphol, P. &Soonthornphisaj, N., Data Mining of Dengue Infection Using Decision Tree, in Latest Advances in Information Science and Applications, pp. 154-159, 2012.

[17]. Boukenze, Basma, HajarMousannif, and AbdelkrimHaqiq. "Predictive analytics in healthcare system using data mining techniques." Comput Sci InfTechnol 1 (2016): 1-9.

[18]. Yang, C.S., Wei, C.P., Yuan, C.C. &Schoung, J.Y., Predicting The Length of Hospital Stay of Burn Patients: Comparison of Prediction Accuracy among Different Clinical Stages, Decision Support Systems, 50, pp. 325- 335, 2010.

[19]. Tomar, D. & Agarwal, S., A Survey on Data Mining Approaches for Healthcare, International Journal of Bio-Science and Bio-Technology, 5, pp. 241-266, 2013.

[20]. Agrawal, Gaurav L., and Hitesh Gupta. "Optimization of C4. 5 decision tree algorithm for data mining application." International Journal of Emerging Technology and Advanced Engineering 3, no. 3 (2013): 341-345.