International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 08 | Aug 2019 www.irjet.net p-ISSN: 2395-0072

Hospital Admission Prediction Model

Twinkle L. Chauhan[1], Mukta S. Takalikar[2], Anil kumar Gupta[3]

[1,2]Dept. of Computer Engineering, PICT, Pune, Maharashtra, India.
[3]CDAC HPC I&E, Pune, Maharashtra, India.

Abstract - *Crowding in Emergency Departments (ED's) has a great impact on the patients. Overcrowding leads to negative consequences in ED. Therefore, it is a need to find innovative methods to improve patient flow and prevent overcrowding. This can be done with the help of different machine learning and deep learning techniques to predict the ED admissions. Predicting the ED admissions will help in resource planning of the hospital management. The proposed model will help to reduce the overall time required in the triage process.*

Key Words: Data Mining, Health Care, Machine Learning.

1. INTRODUCTION

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[1]. 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.

Patients waiting in the EDs lead to crowding and this have negative impacts on management, patients, etc. Therefore, it is a need to explore methods which will help to improve patient flow for preventing overcrowding and reducing the waiting time involved in triage process. It will help to reduce the crowding and also help in resource management. This can be done with the help of different machine learning techniques by predicting the patient's admission.

2. Literature Survey

Chauhan Twinkle[1] and her team has survey different techniques used in the hospital admission prediction. Byron Graham. [2] developed a prediction model in which machine learning techniques such as Logistic Regression, Decision Tree and Gradient Boosted Machine were used. The most important predictors in there model were age, arrival mode, triage category, care group, admission in past-month, past-year. In which the gradient boosted machine outperforms and focus on avoiding the bottleneck in patient flow. Jacinta Lucke. [3] and team has designed the predictive model by considering age as main attribute, where the age is categorized in two category below 70 years and above 70 years. They observed that the category of people below 70 years was less admitted in compare with the category of people above 70 years. Younger patient had higher accuracy while the older patient had high risk of getting admitted to hospital. The decision of prediction was based on the attributes such as age, sex, triage category, mode of arrival, chief complaint, ED revisits, etc. Xingyu Zhang [4] in there predictive model, they have used logistic regression and multilayer neural network. This methods were implemented using the natural language processing and without natural language processing. The accuracy of model with natural language processing is more than the model without natural language processing.

Boukenze. [5] with his team created a model using decision tree C4.5 for predicting admissions which overall gave a good accuracy and less execution time. The author has used the prediction model for predicting a particular disease that is chromic kidney disease. Dinh and his team [6], developed a model which uses multivariable logistic regression for prediction. For the prediction the two main attributes were demographics and triage process, which helped to increase the accuracy. Davood. [7] developed a model for reducing emergency department boarding using the logistic regression and neural network, were a set of thumb rules were developed to predict the hospital admissions. The prediction model used as decision support tool and helped to reduce emergency department boarding. The set of thumb rules were found by examining the importance of eight demographic and clinical factors such as encounter reason, age, radiology exam type, etc. of the emergency department patient's admission. Xie. [8] and his teams model consist of coxian phase type distribution (PH Model) and logistic regression where the PH model has out performs than logistic regression.

Peck and his teams [9] created a model for predicting the inpatient for same-day to improve patient flow. The model uses Naive Bayes and linear regression with logit link function, the result of the model was accurate even though it had less number of independent variables. Sun. [10] and his team uses logistic regression for creating the model with the help of triage process which plays an important role for early prediction of hospital admission.

The factors which were considered for prediction are age, sex, emergency visit in preceding three months, arrival

mode, patient acuity category, coexisting chronic diseases. Jones [11] with his team developed a predictive model for forecasting the daily patient volumes in emergency department. The model uses regression which is actually a time series regression and exponential smoothing where time series regression performs better than linear regression.

3. METHODOLOGY

As the patient arrives in emergency department, the process called triage is carried out. If the arrived patient is critical in condition then two possibilities are there. First, according to the condition if patient can get treated with the medicine or injections then, the patient is treated with that. Second, if the signs and symptoms are not such then patient goes through surgery. In mean time, the relatives of the patient are asked to fill the causality papers where that patient will be assigned the admission number which refers to the admission in the emergency department.

If the symptoms of the patient are not critical but needs to treated as soon as possible then such patients are given waiting time of around 10 to 15 minutes. In that waiting time different tests and scans are carried out of the patient for detail analysis. If the patient is suffering from acute illness then such patients is given waiting for around 30 to 45 minutes. So, in overall triage process each patient has to wait for some time at least due to less number of beds available in emergency department, then a queue is there for different tests and scans. This makes the emergency department crowded.

When a patient arrives through some road accident or anything critical where the identification of that patient can't be recognized. That time the patient is labeled as unknown and a MLC is registered. MLC is Medico Legal Complaint where, a complaint is registered of that patient to the police department for the identification process of that patient. Another case which emergency departments carries is that when certain patient arrives in emergency department and during the start of triage process if patient is declared to be dead, then that patient is directed to the mortuary without admission in emergency department and the process of the death certificate of that patient starts.

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 past 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 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. 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 record are added to the database of hospital patients and triage is done. The model predicts the condition in which the patient was last time admitted to that hospital, which helps in deciding whether the patient might get admitted this time or not. The whole process reduces the time requirement for admitting the patient from emergency department to the hospital as well reduces the waiting time of the patient while triage. This results in less crowding in the emergency department.

3.1. SYSTEM ARCHITECTURE

The system architecture consists of five steps:

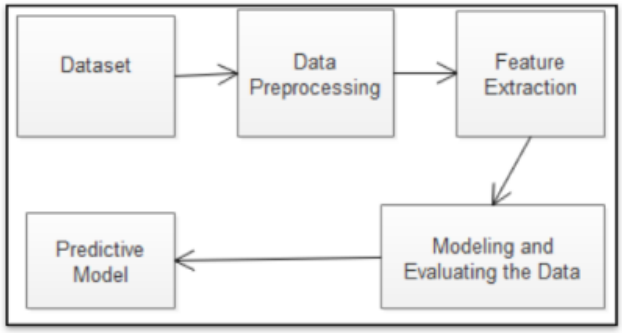

Figure 1: System Architecture

1) Dataset: A hospital dataset is used for the further processing. It is the raw data in the comma separated value (csv) format. It consists of 10 attributes such as ED-level, acuity, etc.
2) Data Preprocessing: The second step is data preprocessing in which all the missing values, null values are removed. Removal of extra value is also done. Format the attributes in correct format.
3) Feature Extraction: A particular number of features are extracted for the model. Such features are selected which are important and which helps in predicting. From the 10 attributes, main attributes are selected. Here five main attributes are selected.
   a) OSHPD-ID: A unique 10 digit number assigned to each patient.
   b) ED-LEVEL: Hospital services providing immediate initial evaluation and treatment to patients on a 24hrs basis.
   c) EMSA-TRAUMA-LEVEL: Emergency Medical
   d) Services Agency trauma center designation level.
   e) ACUITY: Emergency Department type of visit.
   f) ADMISSION-FROM-ED: Total Emergency Department visits by type, resulting in an inpatient admission.
4) Modeling the data: The dataset is divided into two parts, training and testing. In this step, the training dataset is used. Using different machine learning techniques, the model is trained. The trained model obtained in previous step is now used to evaluate. For evaluating the testing dataset is used.

5) Predictive Model: Now after the number of times training and evaluating the model, it is ready for the prediction purpose where external data is given as input.

3.2. MACHINE LEARNING ALGORITHMS AND PERFORMANCE

In this model three machine learning algorithms are used for training purpose: (1) Gradient Boosted Machine, (2) Random Forest, (3) Decision Tree and (4) Extreme Gradient Boosting algorithms along with its variants. Boosting is a class of ensemble learning techniques for classification problems. It aims to build a set of weak learners to create one strong learner. The gradient boosted machine is that algorithm which is a tree based ensemble technique. GBM creates multiple weakly associated decision trees that are combined to get the final prediction. It is also known as boosting model. The second algorithm is the Random forest. This algorithm also uses an ensemble learning approach for classification while training process by creating number of decision trees. The next algorithm is the decision tree which is specifically recursive partitioning. The algorithm splits the data at each node based on the variable that separates the data unless an optimal model is not obtained [2].

XGBoost stands for eXtreme Gradient Boosting. XGBoost is a variant implementation of gradient boosted decision trees designed in such a way to achieve better speed and performance. XGBoost dominates structured or tabular datasets on classification and regression predictive modeling problems. XGBoost algorithm performs linear regression technique to predict the results. Linear regression algorithm is that technique which is used to find linear relationship between target and one or more predictors. Logistic regression is a classification algorithm which is used to classify or predict between categorical variable. In binary logistic regression, the dependent variable is binary in nature. XGBoost with multiclass classification classifies into multiple classes. Here the final predicting classes are more than two that is multiple. When the final predictor has more number of classes that time multiclass classification is used.

The RPART, CARET packages are used for the implementation of the above algorithm. As decision tree works on single tree and the remaining machine learning algorithm works on ensemble of trees therefore, these packages are helpful for the implementation. The CARET package was used to train and tune the machine learning algorithms[1]. This library provides a consistent framework to train and tune models. The performances of the machine learning algorithms are evaluated by the range of measures such as Accuracy, Cohens Kappa, Rooted Mean Square Error.

4. RESULTS AND DISCUSSION

For the evaluation of the methods accuracy, kappa, sensitivity and specificity this performance metrics are used. As shown in Figure 2, a comparison of the accuracy parameter is carried out between the variants of the XGBoost algorithm. XGBoost with multiclass classification has outperformed than the remaining two algorithms. Figure 3, shows comparison of all the algorithms were also, the XGBoost with multiclass classification has outperformed.

The above study gives a broader view of different methods of machine learning used in the field of healthcare. Due to the prediction, the process time required for triage gets reduced. Even the admission process time gets reduced. Along with this it also helps in planning the resources in the emergency department as well in the hospital.

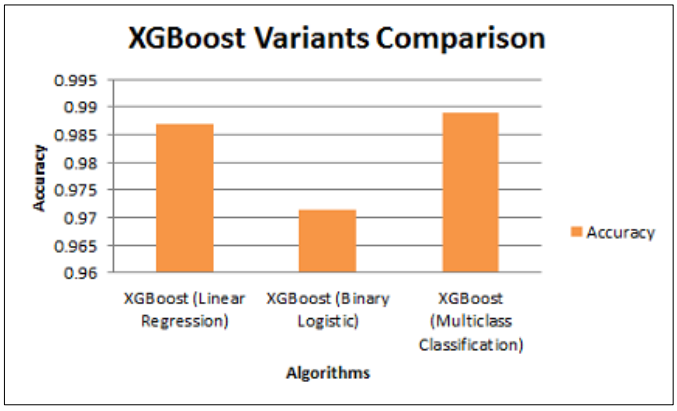

Figure 2: XGBoost Variants Comparison

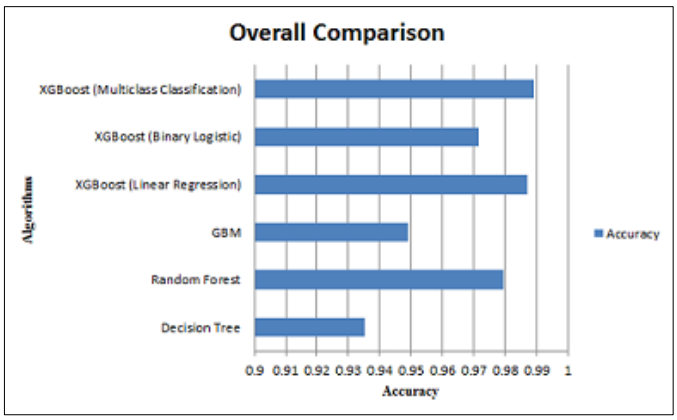

Figure 3: Overall Comparison

5. CONCLUSION AND FUTURE WORK

The study involved a survey of different methods used for the prediction model of hospital admission. Along with this, comparison of different machine learning algorithms namely, decision tree, random forest, gradient boosted machine and extreme gradient boosting techniques is

done which are implemented for the model. Overall the extreme gradient boosting performs better when compared to the rest of techniques. Implementation of these models will help the hospital decision makers for planning and managing the hospital resources based on the patient flow. This would help for reducing the emergency department crowding.

Current system is using the machine learning techniques for predictions, further deep learning techniques can be used. Even in machine learning different algorithms can be ensemble. Along with the hospital textual data, images can also be embedded for more efficiency and accurate results.

REFERENCES

[1] 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.

[2] Graham, Byron and Bond, Raymond and Quinn, Michael and Mulvenna, Maurice," Using Data Mining to Predict Hospital Admissions from the Emergency Department," IEEE Access., vol.6, pp.10458–10469, 2018.

[3] Lucke et al.,"Early prediction of hospital admission for emergency department patients: a comparison between patients younger or older than 70 years," Emerg Med J., vol.35, pp.18–27, 2018.

[4] Zhang, Xingyu and Kim, Joyce and Patzer, Rachel E and Pitts, Stephen R and Patzer, Aaron and Schrager, Justin D, "Prediction of emergency department hospital admission based on natural language processing and neural networks," Methods of information in medicine., vol.56, pp. 377– 389, 2017.

[5] Boukenze, Basma and Mousannif, Hajar and Haqiq, Abdelkrim," Predictive analytics in healthcare system using data mining techniques," Computer Sci Inf Technol., vol. 1, pp. 1–9, 2016.

[6] Dinh et al.,"The Sydney Triage to Admission Risk Tool (START) to predict Emergency Department Disposition: A derivation and internal validation study using retrospective state-wide data from New South Wales, Australia," BMC emergency medicine., vol.16, pp. 46, 2016.

[7] Golmohammadi, Davood," Predicting hospital admissions to reduce emergency department boarding," International Journal of Production Economics., vol. 182, pp. 535–544, 2016.

[8] Xie, Bin," Development and validation of models to predict hospital admission for emergency department patients," International Journal of Statistics in Medical Research., vol.2, pp. 55–66, 2013.

[9] Peck, Jordan S and Benneyan, James C and Nightingale, Deborah J and Gaehde, Stephan A, "Predicting emergency department inpatient admissions to improve same-day patient flow," Academic Emergency Medicine., vol. 19, pp. E1045–E1054, 2012.

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

[11] Jones et al.," Forecasting daily patient volumes in the emergency department," Academic Emergency Medicine., vol. 15, pp. 159–170, 2008.