International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

Using Data mining to predict hospital admissions from the emergency department

V Basha [1],D Shilpa[2], Pendom Poojitha[3],Cheruku Manaswini[4], Yerramshetti Manikanta[5]
[2,3,4,5] UG Scholars, Department of CSE, **AVN Institute of Engineering and Technology**,Hyderabad, Telangana, India.
[1] Assistant Professor, Department of CSE, **AVN Institute of Engineering and Technology**, Hyderabad, Telangana, India.

ABSTRACT

Crowding within emergency departments (EDs) can have significant negative consequences for patients. EDs therefore need to explore the use of innovative methods to improve patient flow and prevent overcrowding. One potential method is the use of data mining using machine learning techniques to predict ED admissions. This paper uses routinely collected administrative data (120 600 records) from two major acute hospitals in Northern Ireland to compare contrasting machine learning algorithms in predicting the risk of admission from the ED.

We use three algorithms to build the predictive models:

1) Logistic regression;
2) Decision trees; and
3) Gradient boosted machines (GBM).

The GBM performed better (accuracy D 80:31%, AUC-ROC D 0:859) than the decision tree (accuracy D 80:06%, AUC-ROC D 0:824) and the logistic regression model (accuracy D 79:94%, AUC-ROC D 0:849). Drawing on logistic regression, we identify several factors related to hospital admissions, including hospital site, age, arrival mode, triage category, care group, previous admission in the past month, and previous admission in the past year. This paper highlights the potential utility of three common machine learning algorithms in predicting patient admissions. Practical implementation of the models developed in this paper in decision support tools would provide a snapshot of predicted admissions from the ED at a given time, allowing for advance resource planning and the avoidance bottlenecks in patient flow, as well as comparison of predicted and actual admission rates. When interpretability is a key consideration, EDs should consider adopting logistic regression models, although GBM's will be useful where accuracy is paramount.

INTRODUCTION

Emergency department(ED) crowding can have serious negative consequences for patients and staff, such as increased wait time, ambulance diversion, reduce dstaff morale, adverse patient outcomes such as increased mortality, and cancellation of elective procedures. Previous research has shown ED crowding to be a significant international problem, making it crucial that innovative steps are taken to address the problem. There are arrange of possible causes of ED crowding depending on the context, with some of the main reasons including increased ED attendances, inappropriate attendances, a lack of alternative treatment options, alacko

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

fin patient beds, ED staffing shortages, and closure of other local ED departments. The most significant of these causes is the in ability to transfer patients to an inpatient bed, making it critical for hospitals to manage patient flow and understand capacity and demand for inpatient beds. One mechanism that could help to reduce ED crowding and improve patient flow is the use of data mining to identify patients at high risk of an inpatient admission, therefore allowing measures to be taken to avoid bottlenecks in the system. For example, a model that can accurately predict hospital admissions could be used for inpatient bed management, staff planning and to facilitate specialised work streams within the ED. Cameron also propose that the implementation of the system could help to improve patient satisfaction by providing the patient with advance notice that admission is likely. Such a model could be developed using data mining techniques, which involves examining and analysing data to extract useful information and knowledge on which decisions can be taken. This typically involves describing and identifying patterns in data and making predictions based on past patterns. This study focuses on the use of machine learning algorithms to develop models top redict hospital admissions from the emergency department, and the comparison of the performance of different approaches to model development. We trained and tested the models using data from the administrative systems of two acute hospitals in Northern Ireland. The performance of EDs has been a particular issue for the Northern Ireland healthcare sector in recent years. ED sin Northern Ireland have been facing pressure from an increase in demand which has been accompanied by adverse levels of performance across the region compared to some other areas of the UK. For example, in June 2015 only one Northern Ireland ED department met the 4 hour wait time target, with over 200 patients across the region waiting over 12 hours to be admitted or sent home. This can have a negative impact on patients at various stages of their journey, as presented in high profile incidents reported by the media. Patients attending the ED typically go through several stages between the time of arrival and discharge depending on decisions made at preceding stages. ED attenders can arrive either via the main reception area or in an ambulance. At this point, the patient’s details are recorded on the main ED administration system, before the patient is either admitted, as in severe cases, or proceeds to the waiting area. The patient then waits for a target time of less than fifteen minutes before triage by a specialist nurse. The Manchester Triage scale is used by all Northern Ireland hospitals, and involves prioritising patients based on the severity of their condition, and to identify patients who are likely to deteriorate if not seen urgently and those who can safely wait to be seen. Triage is an important stage in the patient journey to ensure the best use of resources, patient satisfaction, and safety. Triage systems have also been found to be reliable in predicting admission to hospital, but are most reliable at extreme points of the scale, and less reliable for the majority of patients who fall in the mid points. Once triaged, the patient returns to the waiting room, before assessment by a clinician, who will make a recommendation on the best course of action, which could include treatment, admission,

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
ISSN: 2457-0362

follow up at an outpatient clinic or discharge. If there is a decision to admit the patient, the ED sends a bed request to the ward, and the patient continues to wait until the bed is available. Bottle necks or excess demand at any point in this process can result in ED overcrowding. Routine recoding of data on hospital administrative systems takes place at each stage of this process, providing an opportunity to use machine learning to predict future stage sin the process, and in particular, whether there is an admission. This study draws on this data to achieve two objectives. The first is to create a model that accurately predicts admission to hospital from the ED department, and the second is to evaluate the performance of common machine learning algorithms in predicting hospital admissions. We also suggest use cases for the implementation of the model as a decision support and performance management tool.

LITERATURAL SURVEY

Access block causes emergency department overcrowding and ambulance diversion in Perth

Providing acutely ill patients with rapid access to emergency care is the prime role of emergency medicine. Access block refers to the situation where patients in the emergency department (ED) requiring inpatient care are unable to gain access to appropriate hospital beds within a reasonable time frame, [1] resulting in ED overcrowding and ambulance diversion.Overcrowding in the ED has been described as "the most serious issue confronting EDs in the developed world".[2] Overcrowding refers to the situation where ED function is impeded, primarily because the number of patients waiting to be seen, undergoing assessment and treatment, or waiting for departure, exceeds the physical or staffing capacity of the ED.[1]The effects of overcrowding have been previously reported, and include ambulance diversion, impaired access to emergency care, compromised clinical care, prolonged pain and suffering, and prolonged inpatient length of stay, and has been linked to fatalities.[2–7] Ambulance diversion was rare in Perth for most of the 1990s but is now a daily event. All publicly administered emergency departments in metropolitan Perth use the Emergency Department Information System (EDIS), with continuous electronic capture of ED patient transit through the ED. As the most isolated capital city in the world, there is virtually no leakage of ED patients from Perth to other cities. The combination of excellent information and geographical isolation present a unique opportunity to systematically evaluate the possible causes of metropolitan ED overcrowding and ambulance diversion.

Identifying high-risk patients for triage and resource allocation in the ED

Authors: Ruger JP[1], Lewis LM, Richter CJ.

Five-point triage assessment scales currently used in many emergency departments (EDs) across the country have been shown to be accurate and reliable. We have found the system to be highly predictive of outcome (hospital admission, intensive care unit/operating room admission, or death) at either extreme of the scale but much less predictive in the middle triage group. This is problematic because the middle triage acuity group is the largest, in our experience comprising almost half of all patients. Patients triaged to the 2 highest acuity categories (A and B) have admission/ED death rates of 76% and 43%, respectively. In contrast, the 2 lowest acuity categories (D and E) have admission/ED death rates of 1% or less. The middle category (C), however, has an overall admission/ED death rate of 10%, too high to be comfortable with prolonged delays in the ED evaluation of these patients. We studied this group to determine if easily obtainable clinical factors could identify higher-risk patients in this heterogeneous category. Data were obtained from a retrospective, cross-sectional study of all patients seen in 2001 at an urban academic hospital ED. The main outcome measure for multivariate logistic regression models was hospital admission among patients triaged as acuity C. Acuity C patients who were 65 years or older, presenting with weakness or dizziness, shortness of breath, abdominal pain, or a final diagnosis related group diagnosis of psychosis, were more likely to be admitted than patients originally triaged in category B. These findings suggest that a few easily obtainable clinical factors may significantly improve the accuracy of triage and resource allocation among patients assigned with a middle-acuity score.

Applied predictive modelling

Every day people are faced with questions such as “What route should I take to work today?” “Should I switch to a different cell phone carrier?” “How should I invest my money?” or “Will I get cancer?” These questions indicate our desire to know future events, and we earnestly want to make the best decisions towards that future. We usually make decisions based on information. In some cases we have tangible, objective data, such as the morning traffic or weather report. Other times we use intuition and experience like “I should avoid the bridge this morning because it usually gets bogged down when it snows” or “I should have a PSA test because my father got prostate cancer.” In either case, we are predicting future events given the information and experience we currently have, and we are making decisions based on those predictions. As information has become more readily available via the internet and media, our desire to use this information to help us make decisions has intensified. And while the human brain can consciously and subconsciously assemble a vast amount of data, it cannot process the even greater amount of easily obtainable, relevant information for the problem at hand. To aid in our decision-making processes, we now turn to tools like Google to filter billions of web pages to find the most appropriate information for our queries, WebMD to diagnose our illnesses based on our symptoms, and E*TRADE to screen thousands of stocks and identify the best investments for our portfolios. These sites, as well as many others, use tools that take our current information, sift through data looking for patterns that are relevant to our problem, and return answers. The process of developing these kinds of tools has evolved throughout a number of fields such as chemistry, computer science, physics, and statistics and has been called “machine learning,”“artificial intelligence,” “pattern recognition,” “data mining,” “predictive analytics,” and “knowledge discovery.” While each field approaches the problem using different perspectives and tool sets, the ultimate objective is the same: to make an accurate prediction. For this book, we will

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
IJARST
ISSN: 2457-0362

pool these terms into the commonly used phrase predictive modeling

SYSTEM ANALYSIS

Existing System:

- Sun developed a logistic regression model using two years of routinely collected administrative data to predict the probability of admission at the point of triage. Risk of admission was related to age, ethnicity, arrival mode, patient acuity score, existing chronic conditions, and prior ED attendances or admission in the past three months. Although their data showed the admission of more females than males, sex was not significant in the final model.

- Similarly, Cameron developed a logistic regression model to predict the probability of admissions at triage, using two years of routine administration data collected from hospitals in Glasgow. The most important predictors in their model included `triage category, age, National Early Warning Score, arrival by ambulance, referral source, and admission within the last year', with an area under the curve of the receiver operating characteristic (AUC-ROC) of 0.877. Kim used routine administrative data to predict emergency admissions, also using a logistic regression model. However, their model was less accurate with an accuracy of 76% for their best model.

- Although these models highlight the usefulness of logistic regression in predicting ED admissions, Xie achieved better performance using a Coxian Phase model over logistic regression model, with the former AUC-ROC of 0.89, and the latter 0.83.Wang used a range of machine learning algorithms to predict admissions from the ED, comparing the ability of fuzzy min-max neural networks (FMM) to other standard data mining algorithms including classication and regression trees (CART), Multi Layer Perceptron (MLP), random forest, and AdaBoost.

- Similarly, Peck developed three models to predict ED admissions using logistic regression models, naïve Bayes, and expert opinion. All three techniques were useful in predicting ED admissions. Variables in the model included age, arrival mode, emergency severity index, designation, primary complaint, and ED provider. Their logistic regression model was the most accurate in predicting ED admissions, with an AUC-ROC of 0.887. Perhaps surprisingly, this model performed better than triage nurse's opinion regarding likely admission. The use of logistic regression to predict admission was subsequently found to be generalizable to other hospitals. Using simulation models, Peck have shown that the use of the predictive models to prioritise discharge or treatment of patients can reduce the amount of time the patient spends in the ED department.

Disadvantages:

- There are no Methods for Predicting admissions based on the test data set.
- There is no Data Set extraction using fast techniques.

Proposed System:

- ❖ The proposed system implemented to reduce ED crowding and improve patient how is the use of data mining to identify patients at high risk of an inpatient admission, therefore allowing measures to be taken to avoid bottlenecks in the system. For example, a model that can accurately predict hospital admissions could be used for inpatient bed management, staff planning and to facilitate specialized work streams within the ED.
- ❖ Cameron also propose that the implementation of the system could help to improve patient satisfaction by providing the patient with advance notice that admission is likely. Such a model could be developed using data mining techniques, which involves examining and analyzing data to extract useful information and knowledge on which decisions can be taken. This typically involves describing and identifying patterns in data and making predictions based on past patterns. This study focuses on the use of machine learning algorithms to develop models to predict hospital admissions from the emergency department, and the comparison of the performance of different approaches to model development. We trained and tested the models using data from the administrative systems of two acute hospitals in Northern Ireland.

Advantages:

- Huge Data Set extraction using fast techniques.
- Data cleansing and feature engineering.
- Data visualization and descriptive statistics.
- Data splitting into training (80%) and test sets (20%) in an effective way.
- Model tuning using the training set and 10-fold cross validation repeated 5 times.
- Predicting admissions based on the test data set
- The evaluation of model performance based on predictions made on the test data.

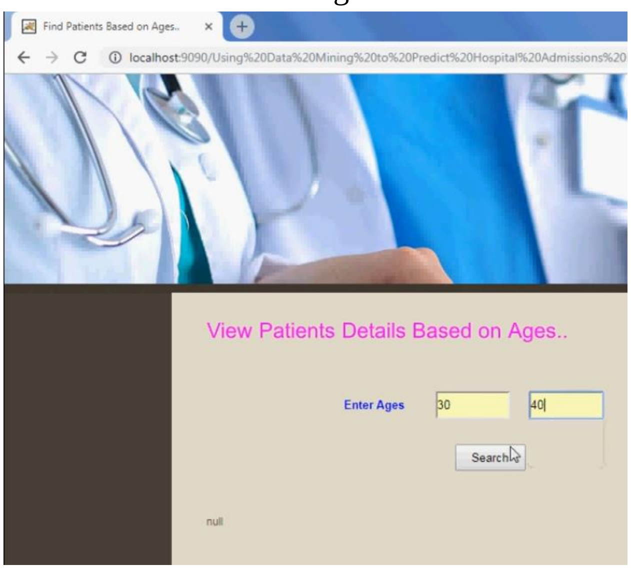
Hospital Admin Home Page

View And Authorize Users

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
IJARST
ISSN: 2457-0362

View And Authorize data Holder

View Patients Batween Ages

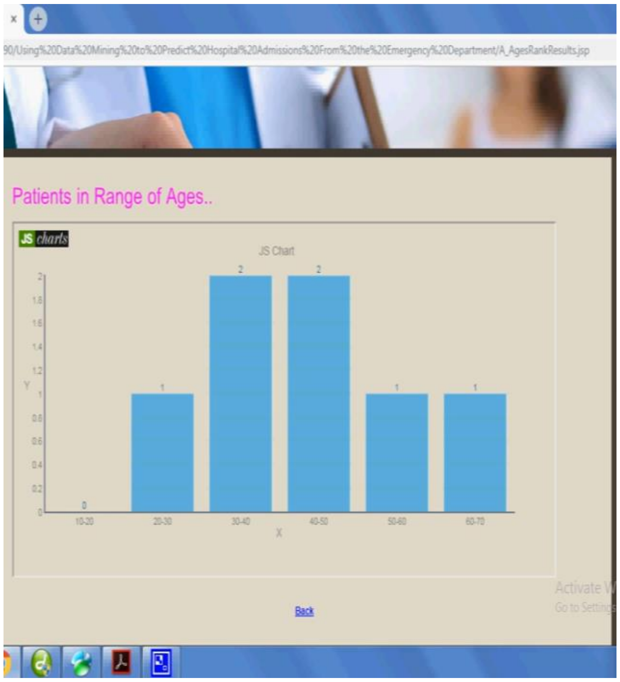


View Patients Search Tranactions

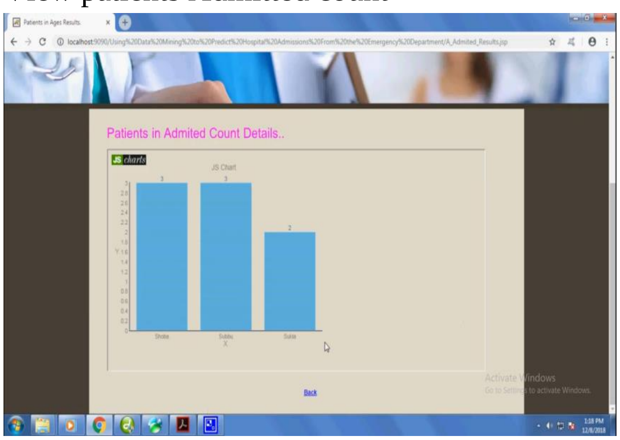
View All Emergency Admitted Patients

View Patients In range of ages

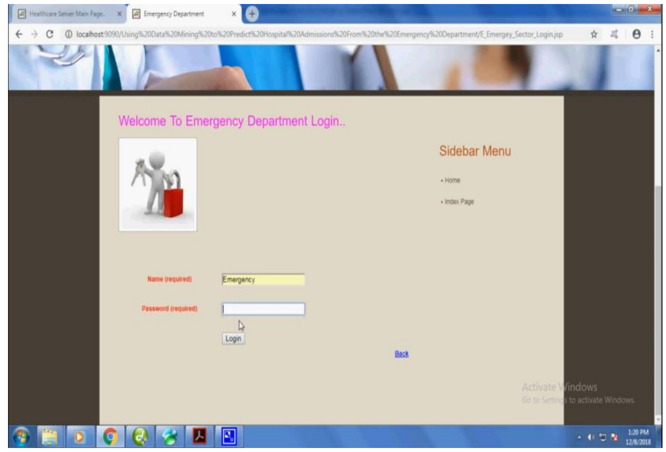
International Journal For Advanced Research
In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362



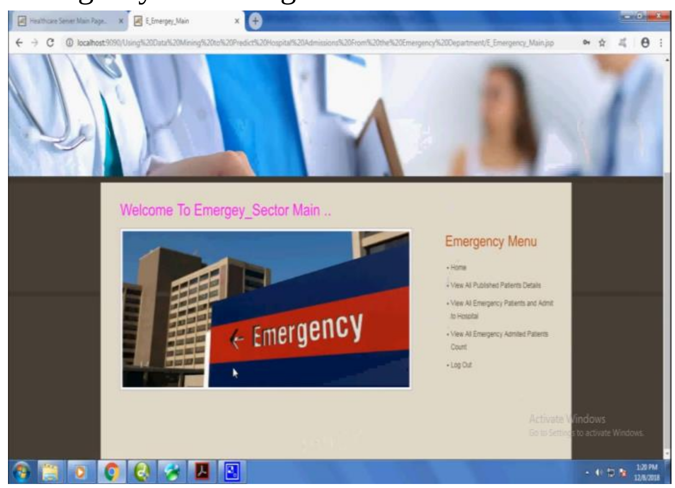
View patients Admitted count



Emergency Department login

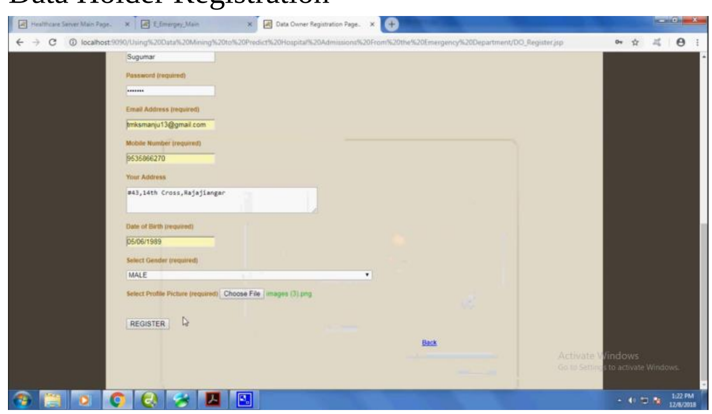


Emergency Home Page



View All patient Details

Data Holder Registration

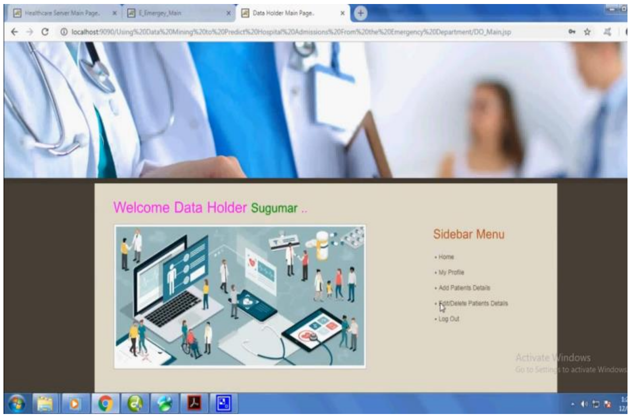


View And Authorize   Data Holders

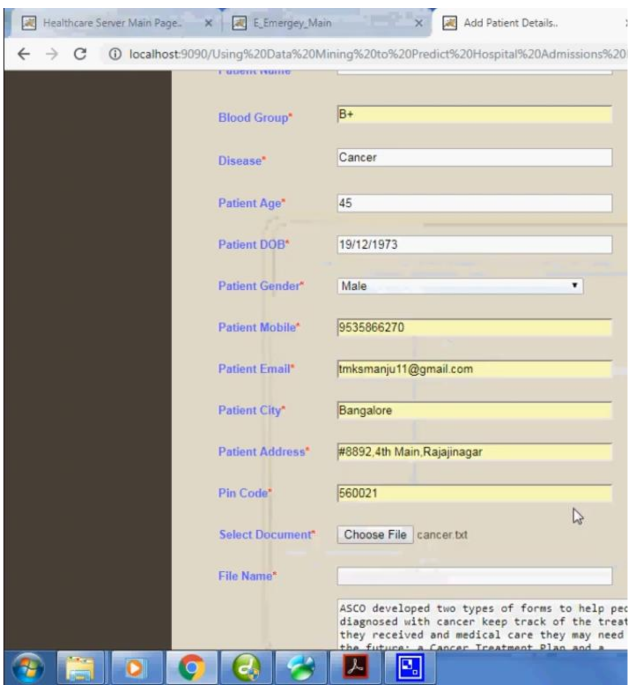
International Journal For Advanced Research In Science & Technology
A peer reviewed international journal www.ijarst.in
IJARST
ISSN: 2457-0362

Data Holder Login

Data Holder Home Page

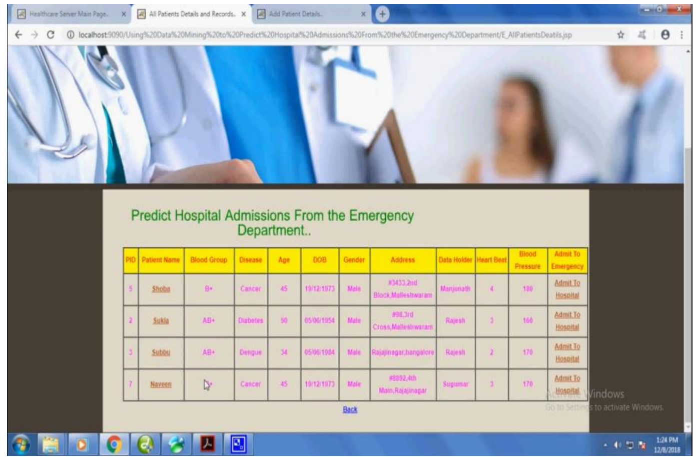


Add Patient Details

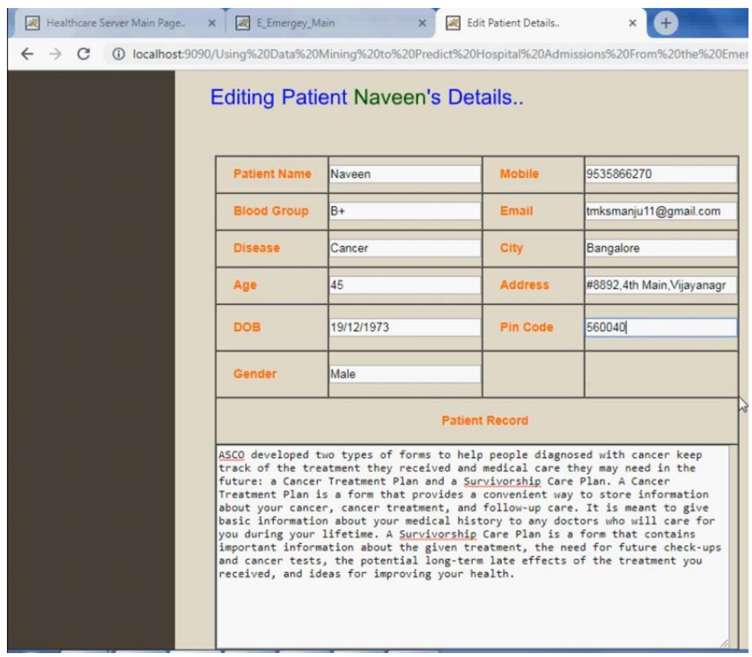


Patient Details Uploaded

View In Emergency



Edit Patient Details



CONCLUSION

This study involved the development and comparison of three machine learning models aimed at predicting hospital admissions from the ED. Each model was trained using routinely collected ED data using three different data mining algorithms, namely logistic regression, decision trees and gradient boosted machines. Overall, the GBM performed the best when compared to logistic regression and decision trees, but the decision tree and logistic regression also performed well. The three models presented in this study yield comparable, and in some cases improved performance compared to models presented in other studies. Implementation of the models as a decision support tool could help hospital decision makers to more effectively plan and manage resources based on the expected patient in flow from the ED. This could help to improve patient flow and reduce ED crowding, therefore reducing the adverse effects of ED crowding and improving patient satisfaction. The models also have potential application in performance monitoring and audit by comparing predicted admissions against actual admissions. However, whilst the model could be used to support planning and decision making, individual level admission decisions still require clinical judgement.

REFERENCES

[1] J. S. Olshaker and N. K. Rathlev, ''Emergency department overcrowding and ambulance diversion: The impact and potential solutions of extended boarding of admitted patients in the emergency department,'' J. Emerg. Med., vol. 30, pp. 351–356, Apr. 2006, doi: 10.1016/j.jemermed.2005.05.023.

[2] J. Boyle et al., ''Predicting emergency department admissions,'' Emerg. Med. J., vol. 29, pp. 358–365, May 2012, doi: 10.1136/emj.2010.103531.

[3] S.L.Bernstein et al., ''The effect of emergency department crowding on clinicallyorientedoutcomes,''Acad.Emerg.Med.,vol.16,no.1,pp.1–10, 2009, doi: 10.1111/j.1553-2712.2008.00295.x.

[4] D. M. Fatovich, Y. Nagree, and P. Sprivulis, ''Access block causes emergency department overcrowding and ambulance diversion in Perth, Western Australia,'' Emerg. Med. J., vol. 22, no. 5, pp. 351–354, 2005, doi: 10.1136/emj.2004.018002.

[5] M. L. McCarthy et al., ''Crowding delays treatment and lengthens emergency department length of stay, even among high-acuity patients,'' Ann. Emerg. Med., vol. 54, no. 4, pp. 492–503, 2009, doi: 10.1016/j.annemergmed.2009.03.006.

[6] D. B. Richardson, ''Increase in patient mortality at 10 days associated with emergency department overcrowding,'' Med. J. Aust., vol. 184, no. 5, pp. 213–216, 2006.

[7] N.R.Hootand D.Aronsky, ''Systematic review of emergency department crowding: Causes, effects, and solutions,''Ann.Emerg.Med.,vol.52,no.2, pp. 126–136, 2008, doi: 10.1016/j.annemergmed.2008.03.014.

[8] Y.Sun, B.H.Heng, S.Y.Tay, and E.Seow, ''Predicting hospital admissions at emergency department triage using routine administrative data,'' Acad. Emerg. Med., vol. 18, no. 8, pp. 844–850, 2011, doi: 10.1111/j.15532712.2011.01125.x.