Quest Journals
Journal of Software Engineering and Simulation
Volume 8 ~ Issue 7 (2022) pp: 09-18
ISSN(Online) :2321-3795 ISSN (Print):2321-3809
www.questjournals.org

Research Paper

Using Data Mining to Predict Hospital Admissions from the Emergency Department

Sanpathi Indraja
Master of Computer Applications
Miracle Educational Society Group of Institutions Institutions
Bhogapuram – Vizianagram-(AP)

Saragadam Sridhar
Master of Computer Applications
Miracle Educational Society Group of
Bhogapuram– Vizianagram-(AP)

ABSTRACT
Crowding within emergency departments 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.
Decision support systems may be able to offer a picture of expected ED admissions at any given moment as a consequence of this study, allowing for resource planning ahead of time and avoiding patient flow bottlenecks. This research also suggests that the models described in this study may be utilised to perform comparisons between projected and actual admission rates. Generalised bivariate models (GBMs) are sufficient when interpretability is a concern; however, if accuracy is crucial, logistic regression models should be considered. INDEX TERMS{IT} Datamining,emergency department, hospitals, machine learning, and predictive models.

Keywords: ED (Emergency Department),GBM (Gneralized bivariate Model, IT (Index Terms).

***R**eceived 25 June, 2022; **R**evised 05 July, 2022; **A**ccepted 07 July, 2022 © The author(s) 2022.*
P**ublished with open access at **www.questjournals.org

I. INTRODUCTION

Emergency department (ED) crowding can have serious negative consequences for patients and staff, such as increased wait time, ambulance diversion, reduced staff 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, a lack of inpatient 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, 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 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 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 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. The performance of EDs has been a particular issue for the Northern Ireland healthcare sector in recent years. EDs in 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 attendees can arrive either via the main reception area or in an ambulance.

II. LITERATURE SURVEY

[1] Byron Graham 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 their model were age, arrival mode, triage category, care group, admission in past-month, past-year. In which the gradient boosted machine out performs and focus on avoiding the bottleneck in patient flow.

[2] Jacinta Lucke and team has designed the predictive model by considering age as main attribute, where the age is categorized in two categories below 70 years and above70 years. They observed that the category of people below 70 years was less admitted when compared with the category of people above 70 years. Younger patient group had higher accuracy while the older patient group 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.

[3] Xingyu Zhang in their predictive model, they have used logistic regression and multilayer neural network. These methods were implemented using natural language processing and without using natural language processing. The accuracy of model with natural language processing is more than the model without natural language processing.

[4]Boukenze 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 chronic kidney disease.

[5] Dinh and his team 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.

III. PROPOSED METHOD

The method for this study involved seven data mining tasks. These were: 1. Data extraction; 2. Data cleansing and feature engineering; 3. Data visualisation and descriptive statistics; 4. Data splitting into training (80%) and test sets (20%); 5. Model tuning using the training set and 10-fold cross validation repeated 5 times; 6. predicting admissions based on the test data set and; 7. The evaluation of model performance based on predictions made on the test data. These steps help to ensure the models are optimal and prevent against overfitting. The study was based on administrative data, all of which was recorded on electronic systems, and subsequently warehoused for business intelligence, analytics, and reporting purposes. The data was recorded during the 2015 calendar year, and includes all ED attendances at two major acute hospitals situated within a single Northern Ireland health and social care trust.

The trust itself offers a full range of acute, community, and social care services delivered in a range of settings including two major acute hospitals, which were the setting for this study. Both hospitals offer a full range of inpatient, outpatient, and emergency services and have close links to other areas of the healthcare system such as community and social services. Hospital 1 is larger, treating approximately 60000 inpatients and day cases each year and 75000 outpatients, whilst hospital 2 treats approximately 20000 inpatients and day cases and 50000 outpatients. The data used in the model building was recorded on the main administrative computer system at each stage of the patient journey at the time the event occurs. A range of variables were considered in the model building, with the final variables decided upon based on previous studies, significance in the models, and the impact of inclusion on the performance of the model. The final models consisted of variables describing whether the patient was admitted to hospital; hospital site; date and time of attendance; age; gender; arrival model; care group; Manchester triage category; and whether the patient had a previous admission to the hospital within the last week, month, or year. The care group is a series of categories indicating the pathway a patient should take. The Manchester triage category is a scale rating the severity of the condition, and used for prioritisation. Prior admissions were measured objectively by querying the hospital database. Feature engineering was also carried out on the date of attendance to disaggregate it into components relating to year, day of the week, and month of the year. The dependent variable in all models was admission to the hospital from the ED. Most of the variables included in the model are mandatory on the ED system, and recorded using of drop down menus. This led to a relatively clean dataset for analysis, with list wise deletion of cases with missing data. Patients attending direct assessment units and observation units are excluded from the analysis, as these patients follow a different pathway to those attending the main ED. Furthermore, many hospitals do not have such departments, which would limit the generalizability of the results. The final dataset consisted of 120,600 observations, of which 10.8% had missing data, leaving 107,545 cases for building the models. To enable validation of the model, random stratified sampling was used to split the data into training (80% of cases) and test (20% of cases) datasets. Data was extracted and stored using SQL Server (2012), and the machine learning and exploratory analysis was carried out using the R software for statistical computing, version 3.2.1.

Random Forest Algorithm

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of ensemble learning, which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model.

As the name suggests, "Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset." Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output.

The greater number of trees in the forest leads to higher accuracy and prevents the problem of overfitting.

The below diagram explains the working of the Random Forest algorithm:

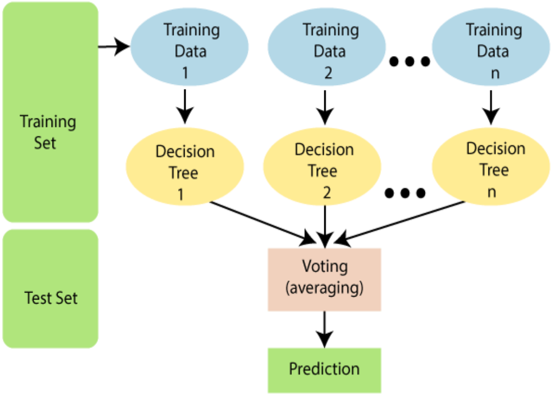


Figure-1: Random Forest Algorithm

Random Forest Advantages

Below are some points that explain why we should use the Random Forest algorithm:

- It takes less training time as compared to other algorithms.
- It predicts output with high accuracy, even for the large dataset it runs efficiently.
- It can also maintain accuracy when a large proportion of data is missing.

Random Forest Algorithm Work

Random Forest works in two-phase first is to create the random forest by combining N decision tree, and second is to make predictions for each tree created in the first phase.

The Working process can be explained in the below steps and diagram:

Step-1: Select random K data points from the training set.

Step-2: Build the decision trees associated with the selected data points (Subsets).

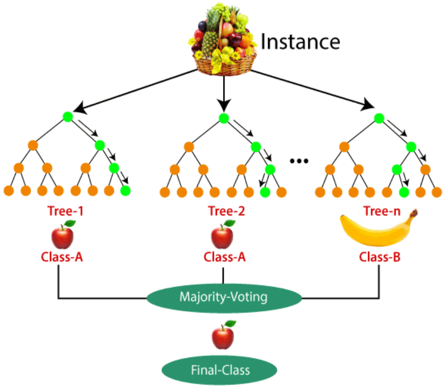
Step-3: Choose the number N for decision trees that you want to build.

Step-4: Repeat Step 1 & 2.

Step-5: For new data points, find the predictions of each decision tree, and assign the new data points to the category that wins the majority votes.

The working of the algorithm can be better understood by the below example:

Example: Suppose there is a dataset that contains multiple fruit images. So, this dataset is given to the Random forest classifier. The dataset is divided into subsets and given to each decision tree. During the training phase, each decision tree produces a prediction result, and when a new data point occurs, then based on the majority of results, the Random Forest classifier predicts the final decision. Consider the below image:



Gradient boosting Algorithm

Boosting is one of the popular learning ensemble modelling techniques used to build strong classifiers from various weak classifiers. It starts with building a primary model from available training data sets then it identifies the errors present in the base model. After identifying the error, a secondary model is built, and further, a third model is introduced in this process. In this way, this process of introducing more models is continued until we get a complete training data set by which model predicts correctly.

AdaBoost (Adaptive boosting) was the first boosting algorithm to combine various weak classifiers into a single strong classifier in the history of machine learning. It primarily focuses to solve classification tasks such as binary classification.

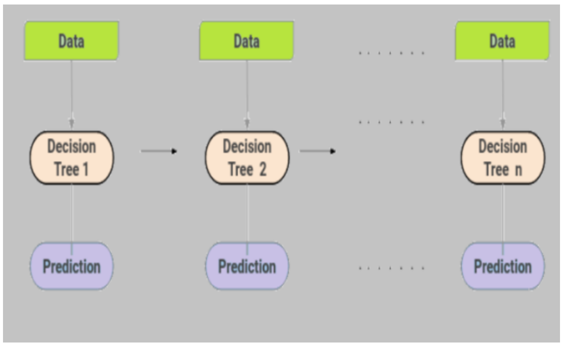


Figure-2: Gradient Boosting Algorithm

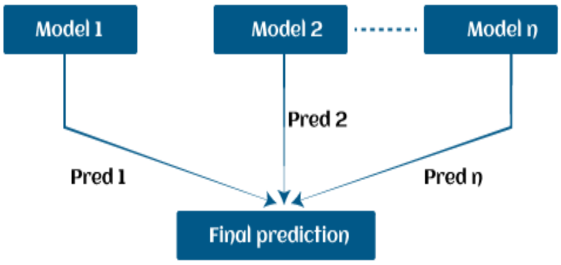
Steps Involved in Boosting Algorithm:

There are a few important steps in boosting the algorithm as follows:

- Consider a dataset having different data points and initialize it.
- Now, give equal weight to each of the data points.
- Assume this weight as an input for the model.
- Identify the data points that are incorrectly classified.
- Increase the weight for data points in step 4.
- If you get appropriate output then terminate this process else follow steps 2 and 3 again.

Example:

Let's suppose, we have three different models with their predictions and they work in completely different ways. For example, the linear regression model shows a linear relationship in data while the decision tree model attempts to capture the non-linearity in the data as shown below image.



Further, instead of using these models separately to predict the outcome if we use them in form of series or combination, then we get a resulting model with correct information than all base models. In other words, instead of using each model's individual prediction, if we use average prediction from these models then we would be able to capture more information from the data. It is referred to as ensemble learning and boosting is also based on ensemble methods in machine learning.

GBM Working procedure

Generally, most supervised learning algorithms are based on a single predictive model such as linear regression, penalized regression model, decision trees, etc. But there are some supervised algorithms in ML that depend on a combination of various models together through the ensemble. In other words, when multiple base models contribute their predictions, an average of all predictions is adapted by boosting algorithms.

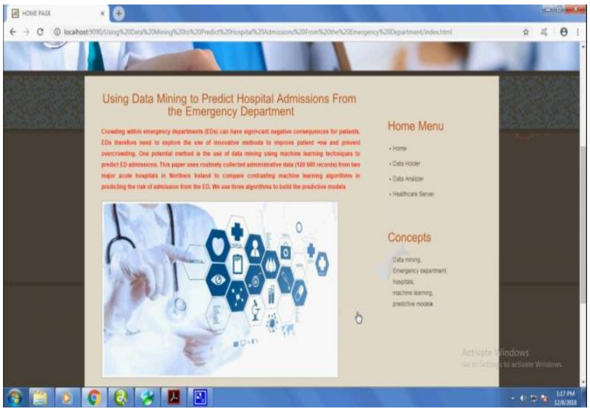
Gradient boosting machines consist 3 elements as follows:

- Loss function
- Weak learners
- Additive model

MACHINE LEARNING ALGORITHMS AND PERFORMANCE

Three machine learning algorithms were applied to the training data to build the models: (1) logistic regression, (2) a decision tree, and (3) gradient boosted machines (GBM). Logistic regression is suitable for predicting a binary dependent variable, such as positive/negative; deceased/alive; or in this study, admit/not admit. The technique uses a logic link function to enable the calculation of the odds of an outcome occurring. The second algorithm that was used was a decision tree, specifically recursive partitioning from the RPART package . The RPART package is an implementation based on the model presented by Breiman and colleagues. This algorithm splits the data at each node based on the variable that best separates the data until either an optimal model is identified or a minimum number of observations exist in the final (terminal) nodes. The resulting tree can then be pruned to prevent over fitting and to obtain the most accurate model for prediction . The third algorithm was a GBM, which creates multiple weakly associated decision trees that are combined to provide the final prediction . recursive partitioning are the complexity parameter and maximum node depth, and for GBM the user can tune the interaction depth, minimum observations in a node, learning rate, and number of iterations. The CARET package was used to train and tune the machine learning algorithms. This library provides the user with a consistent framework to train and tune models, as well as a range of helper functions. To further prevent against over fitting and to evaluate the performance of the models, predictions were made on an unseen test dataset. The performance of each machine learning algorithm was evaluated using a range of measures including accuracy, Cohens Kappa, c-statistics of the ROC, sensitivity and specificity. When interpreting the AUC-ROC, values of between 0.7 and 0.8 can be interpreted as having good discrimination ability, and models with AUC-ROC of greater than 0.8 can be interpreted as having excellent discrimination ability, with values above 0.9 indicating outstanding ability.

IV. RESULTS

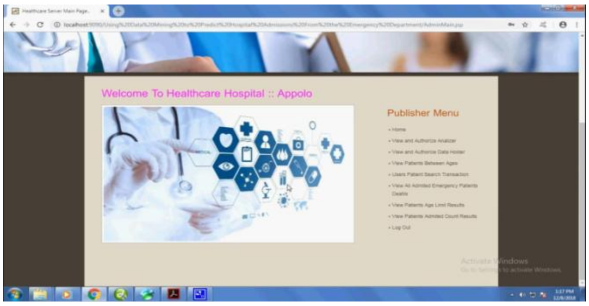


Using Data Mining to Predict Hospital Admissions From the Emergency Department

Crowding within emergency departments (EDs) can have signi+cant negative consequences for patients. EDs therefore need to explore the use of innovative methods to improve patient •ow 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

Figure 3: Home Page

Crowding within emergency departments 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

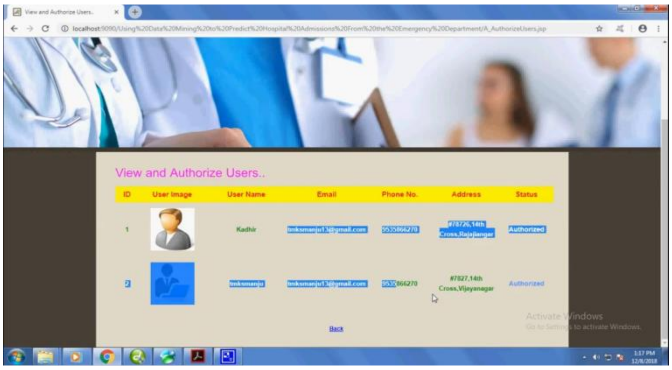


Welcome To Healthcare Hospital :: Appolo

Healthcare Server Main Page..
localhost:9090/Using%20Data%20Mining%20to%20Predict%20Hospital%20Admissions%20From%20the%20Emergency%20Department/AdminMain.jsp
Publisher Menu
• Home
• View and Authorize Analizer
• View and Authorize Data Holder
• View Patients Between Ages
• Users Patient Search Transaction
• View All Admitted Emergency Patients Details
• View Patients Age Limit Results
• View Patients Admitted Count Results
• Log Out
Activate Windows
Go to Settings to activate Windows.
1:17 PM
12/8/2018

Figure 4: Hospital Admin Page

The Health service provider manages a server to provide data storage service and can also do the following operations such as View and Authorize Analyzer, View and Authorize Data Holder, View Patients Between Ages, Users Patient Search Transaction, View All Admitted Emergency Patients Details, View Patients Age Limit Results, View Patients Admitted Count.



View and Authorize Users..

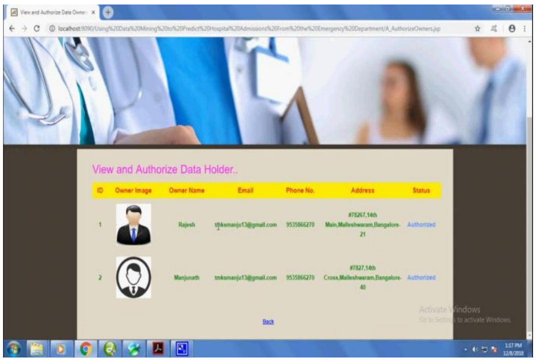
ID	User Image	User Name	Email	Phone No.	Address	Status
1		Kadhir	tmksmanju13@gmail.com	9535866270	#78726,14th Cross,Rajajiangar	Authorized
2		tmksmanju	tmksmanju13@gmail.com	9535866270	#7827,14th Cross,Vijayanagar	Authorized

Back

View and Authorize Users..
localhost:9090/Using%20Data%20Mining%20to%20Predict%20Hospital%20Admissions%20From%20the%20Emergency%20Department/A_AuthorizeUsers.jsp
Activate Windows
Go to Settings to activate Windows.
1:17 PM
12/8/2018

Figure 5: View and Authorize Users

In this output, we can see the details of view and authorize user.



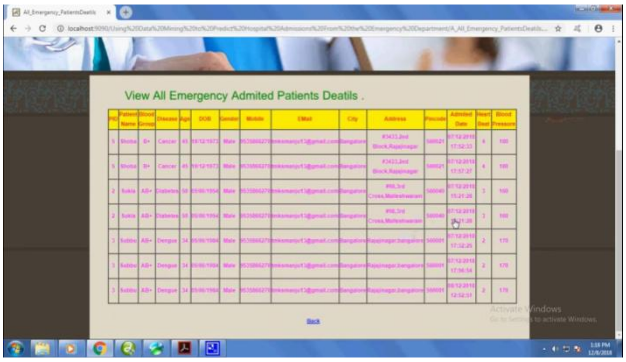
View and Authorize Data Holder..

ID	Owner Image	Owner Name	Email	Phone No.	Address	Status
1		Rajesh	tmksmanju13@gmail.com	9535866270	#78267,14th Main,Malleshwaram,Bangalore-21	Authorized
2		Manjunath	tmksmanju13@gmail.com	9535866270	#7827,14th Cross,Malleshwaram,Bangalore-40	Authorized

Back

Figure 6: View and authorize data holder

In this output, we can see the details of view and authorize Data Holder



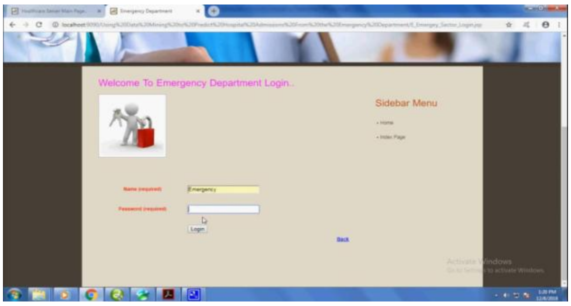
View All Emergency Admited Patients Deatils .

PID	Patient Name	Blood Group	Disease	Age	DOB	Gender	Mobile	EMail	City	Address	Pincode	Admited Date	Heart Beat	Blood Pressure
5	Shoba	B+	Cancer	45	19/12/1973	Male	953586627	tmksmanju13@gmail.com	Bangalore	#3433,2nd Block,Rajajinagar	560021	07/12/2018 17:52:33	4	100
5	Shoba	B+	Cancer	45	19/12/1973	Male	953586627	tmksmanju13@gmail.com	Bangalore	#3433,2nd Block,Rajajinagar	560021	07/12/2018 17:57:27	4	100
2	Sukla	AB+	Diabetes	58	05/06/1954	Male	953586627	tmksmanju13@gmail.com	Bangalore	#98,3rd Cross,Malleshwaram	560040	07/12/2018 15:21:28	3	160
2	Sukla	AB+	Diabetes	58	05/06/1954	Male	953586627	tmksmanju13@gmail.com	Bangalore	#98,3rd Cross,Malleshwaram	560040	07/12/2018 [illegible]	3	160
3	Subbu	AB+	Dengue	34	05/06/1984	Male	953586627	tmksmanju13@gmail.com	Bangalore	Rajajinagar,bangalore	560001	07/12/2018 17:52:25	2	170
3	Subbu	AB+	Dengue	34	05/06/1984	Male	953586627	tmksmanju13@gmail.com	Bangalore	Rajajinagar,bangalore	560001	07/12/2018 17:56:54	2	170
3	Subbu	AB+	Dengue	34	05/06/1984	Male	953586627	tmksmanju13@gmail.com	Bangalore	Rajajinagar,bangalore	560001	08/12/2018 12:52:51	2	170

Back

Figure 7: View all Emergency Admitted Patients Details

In this module, we can see the details of view all emergency admitted patient details.



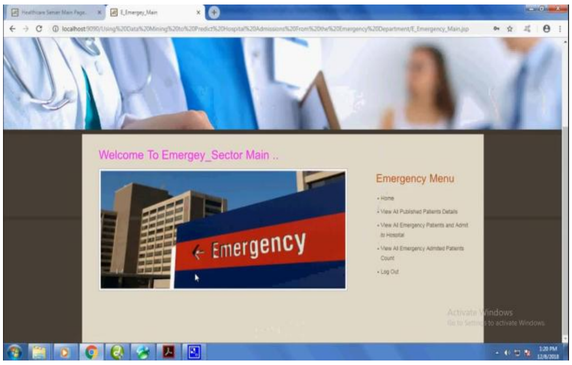
Welcome To Emergency Department Login..

Name (required): Emergency

Password (required):

Figure 8: Emergency Department Login

User need to enter name and password then click on login button .when the user click on home link present in the sidebar menu user need to redirect to the home page .when the user click on index page link present in the sidebar menu user need to redirect to the index page.

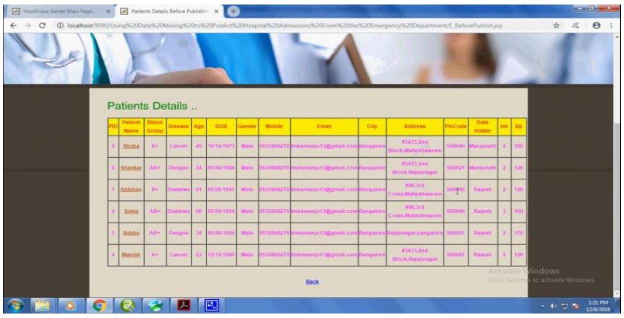


Welcome To Emergey_Sector Main ..

Emergency
Emergency Menu
• Home
• View All Published Patients Details
• View All Emergency Patients and Admit to Hospital
• View All Emergency Admited Patients Count
• Log Out

Figure 9: Emergency Home Page

when the user click on home link present in the emergency menu user need to redirect to the home page .when the user click on View All Published Patients Details link present in the emergency menu user need to redirect to the View All Published Patients Details page when the user click on View All Emergency Patients link present in the emergency menu user need to redirect to the View All Emergency Patients page . when the user click on View All Emergency Admitted Patients Count link present in the emergency menu user need to redirect to the View All Emergency Admitted Patients Count page when the user click on Log Out button then user should logged out.

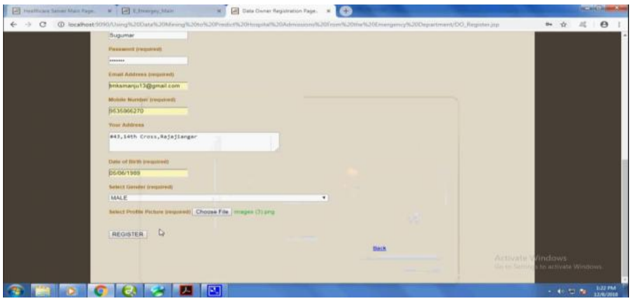


Patients Details ..

PID	Patient Name	Blood Group	Disease	Age	DOB	Gender	Mobile	Email	City	Address	PinCode	Data Holder	Hb	Bp
5	Shoba	B+	Cancer	45	19/12/1973	Male	9535866270	mkmanju13@gmail.com	Bangalore	#3433,2nd Block,Malleshwaram	560040	Manjunath	4	180
6	Shankar	AB+	Dengue	34	05/06/1984	Male	9535866270	mkmanju13@gmail.com	Bangalore	#3433,2nd Block,Rajajinagar	560021	Manjunath	2	120
1	Abhinay	B+	Diabetes	61	05/06/1941	Male	9535866270	mkmanju13@gmail.com	Bangalore	#98,3rd Cross,Malleshwaram	560040	Rajesh	2	120
2	Sukla	AB+	Diabetes	50	05/06/1954	Male	9535866270	mkmanju13@gmail.com	Bangalore	#98,3rd Cross,Malleshwaram	560040	Rajesh	3	180
3	Subbu	AB+	Dengue	34	05/06/1984	Male	9535866270	mkmanju13@gmail.com	Bangalore	Rajajinagar,bangalore	560001	Rajesh	2	170
4	Manish	A+	Cancer	23	12/12/1995	Male	9535866270	mkmanju13@gmail.com	Bangalore	#3433,2nd Block,Rajajinagar	560021	Rajesh	2	120

Figure 10: View All Patients Details

Patient details should be displayed while click on view all published patient details link in the emergency menu



Sugumar

Password (required): ********

Email Address (required): mkmanju13@gmail.com

Mobile Number (required): 9535866270

Your Address: #43,14th Cross,Rajajinagar

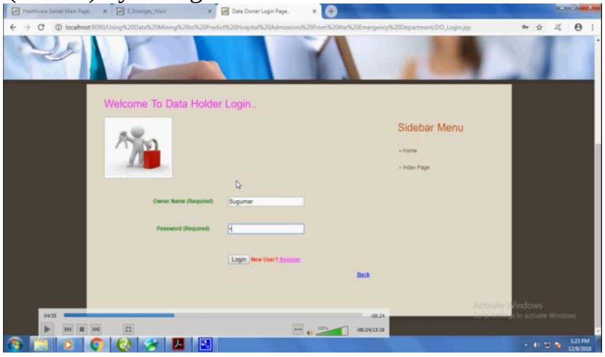
Date of Birth (required): 05/06/1989

Select Gender (required): MALE

Select Profile Picture (required) Choose File images (3).png

Figure 11: Data Holder Registration

Here we can register the user(Owner) by filling the above details



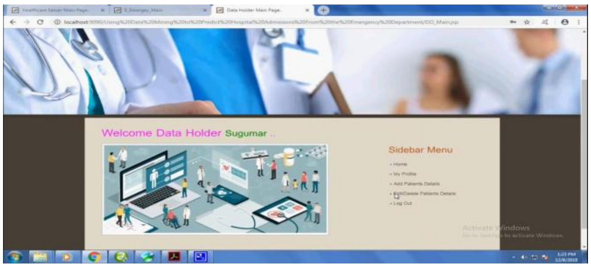
Welcome To Data Holder Login..

Owner Name (Required): Sugumar

Password (Required):

Figure 12: Data Holder Login

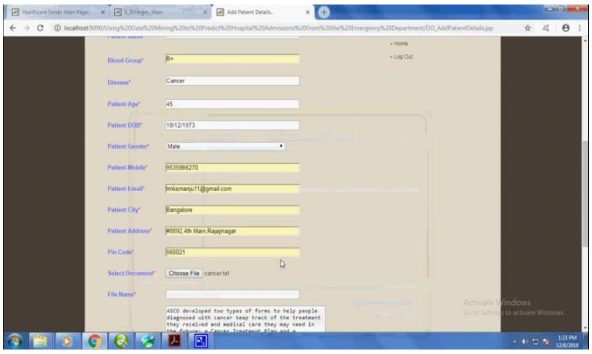
Here we can login with data holder credentials



Welcome Data Holder Sugumar ..

Figure 13: Data Holder Home Page

Here after login with data holder credentials we will redirect to data holder home page. Here we can access data holder profile by clicking the my profile link in the sidebar menu. We can add patient details by clicking the add patient details link in the sidebar menu. We can edit/delete patient details by clicking the edit/delete patient details link in the sidebar menu. User can log out from the data holder home page by clicking logout link in the sidebar menu



localhost:9090/Using%20Data%20Mining%20to%20Predict%20Hospital%20Admissions%20From%20the%20Emergency%20Department/DO_AddPatientDetails.jsp

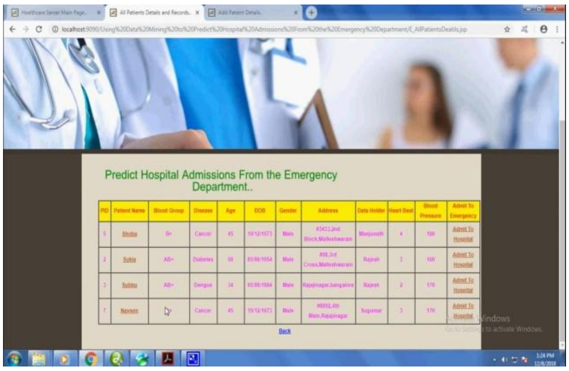
- Home
- Log Out

Blood Group*: B+
Disease*: Cancer
Patient Age*: 45
Patient DOB*: 19/12/1973
Patient Gender*: Male
Patient Mobile*: 9535866270
Patient Email*: tmksmanju11@gmail.com
Patient City*: Bangalore
Patient Address*: #8892,4th Main,Rajajinagar
Pin Code*: 560021
Select Document*: Choose File cancer.txt
File Name*:

ASCO developed two types of forms to help people diagnosed with cancer keep track of the treatment they received and medical care they may need in the future: a Cancer Treatment Plan and a

Figure 14: Add Patient Details

After clicking the add patient details link in the data holder home page it need to redirect to the add patient details page.In the add patient details page we can add the patient details



localhost:9090/Using%20Data%20Mining%20to%20Predict%20Hospital%20Admissions%20From%20the%20Emergency%20Department/E_AllPatientsDetails.jsp

Predict Hospital Admissions From the Emergency Department..

PID	Patient Name	Blood Group	Disease	Age	DOB	Gender	Address	Data Holder	Heart Beat	Blood Pressure	Admit To Emergency
5	Shoba	B+	Cancer	45	19/12/1973	Male	#3433,2nd Block,Malleshwaram	Manjunath	4	180	Admit To Hospital
2	Sukla	AB+	Diabetes	56	05/06/1954	Male	#88,3rd Cross,Malleshwaram	Rajesh	3	160	Admit To Hospital
3	Subbu	AB+	Dengue	34	05/06/1984	Male	Rajajinagar,bangalore	Rajesh	2	170	Admit To Hospital
7	Naveen	[illegible]	Cancer	45	19/12/1973	Male	#8892,4th Main,Rajajinagar	Sugumar	5	170	Admit To Hospital

Back

Figure 15: View in Emergency

In this above output screen, we can see the Admissions from the emergency department

VI.CONCLUSION AND FUTURE WORK

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 reduced 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.

The overall study involved a survey of different methods used for the prediction model of hospital admission. Along with this study it also compares three different machine learning algorithms namely, decision tree, random forest and gradient boosted machine which are used for predicting the hospital admission from the emergency department. Overall the random forest performs better when compared to the decision tree and gradient boosted machine. Implementation of these models could help the hospital decision makers for planning and managing the hospital resources based on the patient flow. This would help reducing the emergency department crowding. In future, different learning and machine learning algorithms can be used to implement the model. Even ensemble of different algorithms can also be done. Different demographics as predictor can be taken into consideration

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 clinically oriented outcomes,’’ 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. Hoot and 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.ann emergmed.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.1553- 2712.2011.01125.x.

[9]. M. A. LaMantia et al., ‘‘Predicting hospital admission and returns to the emergency department for elderly patients,’’ Acad. Emerg. Med., vol. 17, no. 3, pp. 252–259, 2010, doi: 10.1111/j.1553-2712.2009.00675.x.

[10]. J. S. Peck et al., ‘‘Generalizability of a simple approach for predicting hospital admission from an emergency department,’’ Acad. Emerg. Med., vol. 20, pp. 1156–1163, Nov. 2013, doi: 10.1111/acem.12244.

[11]. A. Cameron, K. Rodgers, A. Ireland, R. Jamdar, and G. A. McKay, ‘‘A simple tool to predict admission at the time of triage,’’ Emerg. Med. J., vol. 32, no. 3, pp. 174–179, 2013, doi: 10.1136/emermed-2013-203200.

[12]. N. Esfandiari, M. R. Babavalian, A. M. E. Moghadam, and V. K. Tabar, ‘‘Knowledge discovery in medicine: Current issue and future trend,’’ Expert Syst. Appl., vol. 41, no. 9, pp. 4434–4463, 2014, doi: 10.1016/j.eswa.2014.01.011.

[13]. H. C. Koh and G. Tan, ‘‘Data mining applications in healthcare,’’ J. Healthcare Inf. Manag., vol. 19, no. 2, pp. 64–72, 2005.

[14]. C. Baker. (2015). Accident and Emergency Statistics, England. [Online]. Available: www.parliament.uk/briefing-papers/sn06964.pdf

[15]. DHSSPS. (2015). Northern Ireland Hospital Statistics: Emergency Care (2014/15), Northern Ireland. [Online]. Available: https://www.dhsspsni.gov.uk/sites/default/files/publications/dhssps/hsemergency-care-2014-15.pdf

[16]. M.-L. Connolly. (2015). NI Emergency Healthcare Enquiry Finds ‘Degrading’ Cases. Accessed: Oct. 14, 2015. [Online]. Available: http:// www.bbc.co.uk/news/uk-northern-ireland-32888240

[17]. (2014). Royal Victoria Hospital: Delays ‘Contributed to Five Deaths. Accessed: Oct. 14, 2015. [Online]. Available: http:// www.bbc.co.uk/news/uk-northern-ireland-26135756

[18]. J. P. Ruger, L. M. Lewis, and C. J. Richter, ‘‘Identifying high-risk patients for triage and resource allocation in the ED,’’ Amer. J. Emerg. Med., vol. 25, pp. 794–798, Sep. 2007, doi: 10.1016/j.ajem.2007.01.014. [19] W.-T. Lin, S.-T. Wang, T.-C. Chiang, Y.-X. Shi, W.-Y. Chen, and H.-M. Chen, ‘‘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 Syst. Appl., vol. 37, no. 4, pp. 2733–2741, 2010, doi: 10.1016/j.eswa.2009.08.006.