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

Using Data Mining to Predict Hospital Admissions From the Emergency Department

[1]Dr. P Vijaypal Reddy, [2]Gangadhara Tejasree, [3]Veluguri Navya, [4]Pittala Venkata Sirisha, [5]Kothuri Datta Sainadh
[1,2,3,4,5]Assistant professors, Department of CSE in Narasaraopet Institute Of Technology

ABSTRACT

Crowding within emergency departments (EDs) can have signicant 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: 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 ow, 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.
INDEX TERMS Data mining, emergency department, hospitals, machine learning, predictive models.

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 [1][6]. Previous research has shown ED crowding to be a signicant international problem [7], making it crucial that innovative steps are taken to address the problem [4]. There are a range 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 stafng shortages, and closure of other local ED departments [1], [8]. The most signicant of these causes is the inability to transfer patients to an inpatient bed [1], making it critical for hospitals to manage patient ow and understand capacity and demand for inpatient beds [4]. One mechanism that could help to reduce ED crowding and improve patient ow 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 [9], [10]. 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 [11]. Cameron et al.

[11] 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 [12]. This typically involves describing and identifying patterns in data and making predictions based on past patterns [13]. 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 [14], [15]. 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 [15]. This can have a negative impact on patients at various stages of their journey, as presented in high prole incidents reported by the media [16], [17]. 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 fteen 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 [18]. Triage is an important stage in the patient journey to ensure the best use of resources, patient satisfaction, and safety [19]. 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 [18]. 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, 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. Bottlenecks 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 stages in the process, and in

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

particular, whether there is an admission. This study draws on this data to achieve two objectives. The rst 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.

II. RELATED WORK

Using a range of clinical and demographic data relating to elderly patients, LaMantia et al. [9] used logistic regression to predict admissions to hospital, and ED re-attendance. They predicted admissions with moderate accuracy, but were unable to predict ED re-attendance accurately. The most important factors predicting admission were age, Emergency Severity Index (ESI) triage score, heart rate, diastolic blood pressure, and chief complaint [9] (pg. 255). Baumann and Strout [20] also nd an association between the ESI and admission of patients aged over 65. Boyle et al. [2] used historical data to develop forecast models of ED presentations and admissions. Model performance was evaluated using the mean absolute percentage error (MAPE), with the best attendance model achieving a MAPE of around 7%, and the best admission model achieving a MAPE of around 2% for monthly admissions. The use of historical data by itself to predict future events has the advantage of allowing forecasts further into the future, but has the disadvantage of not incorporating data captured at arrival and through triage, which may improve the accuracy of short term forecasting of admissions. Sun et al. [8] 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 signicant in the nal model. Similarly, Cameron et al. [11] 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 EarlyWarning Score, arrival by ambulance, referral source, and admission within the last year' (pg. 1), with an area under the curve of the receiver operating characteristic (AUC-ROC) of 0.877. Other variables including weekday, out of hours attendances, and female gender, were signicant but did not have high enough odds ratios to be included in the nal models. Kim et al. [21] 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 [22] 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 et al. [23] 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. Overall, MLP and

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

Random Forest 80% of cases correctly, with FMM (with a genetic algorithm) predicting 77.97% of cases correctly. Similarly, Peck et al. [24] 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 [10]. Using simulation models, Peck et al. [25]
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. Qui et al. [26] used a relative vector machine to predict whether an ED attender would be discharged or admitted to one of three hospital words. Their model had an overall
accuracy of 91.9% with an AUC of 0.825. However, the accuracy of predicting the target ward varied by ward and by the probability threshold used. Lucini et al. [27] used eight common machine learning algorithms to predict admissions from the ED department based on features derived from text recorded on the patients record. Six out of the eight algorithms had similar levels of performance including nusupport vector machines, support vector classication, extra trees, logistic regress, random forests, and multinomial naïve bayes, with AdaBoost and a decision tree performing worst. Taking a different approach, Cameron et al. [28] compared the accuracy of nurses predictions of ED admissions with those of an objective score. They nd nurses to be more accurate in cases where they are certain the patient will be admitted, but less accurate than the objective score in cases where they are uncertain about the patient's likelihood of admission.

The literature highlights the application of a range of traditional and machine learning approaches to the prediction
of ED admissions in different contexts using a variety of data. However, there are gaps in the literature to which this
study contributes. Much of the previous work focuses on a narrow range of algorithms, and primarily logistic regression, with fewer studies comparing multiple approaches. This leaves open the potential for the development of more accurate predictive models using other algorithms. For example, gradient boosted machines (GBM) were not applied in any of the studies reviewed, but have been successful in predicting binary outcomes in other scenarios such as hospital transfers and mortality [29]. In addition, few studies were identied that focused on the UK context, and none that focused on Northern Ireland ED's. This is an important gap in the literature as the structure and operation of health services varies considerably between countries and regions within countries. Most previous studies have also tended to focus on developing predictive models for one hospital site, with fewer studies building models using data from multiple sites. This study seeks to contribute to the existing body of knowledge by building machine learning models using a novel dataset and by comparing the performance of less frequently used algorithms with the more traditional logistic regression approach. Moreover, the data used in our study is routinely available at the point of triage,

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

allowing for the potential implementation of a fully automated decision support system based on the models built here.

III. METHODS

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 overtting. 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 nal variables decided upon based on previous studies, signicance in the models, and the impact of inclusion on the performance of the model. The nal 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 engineeringwas 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 listwise 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 nal 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 stratied 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 [32], version 3.2.1. A.

iv 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 logit link function to enable the calculation of the odds of an outcome occurring. The second algorithm that was used was a decision tree, specically recursive partitioning from the RPART package [33]. The RPART package is an implementation based on the model presented by Breiman and colleagues [33], [34]. This algorithm splits the data at each node based on the variable that best separates the data until either an optimal model is identied or a minimum number of observations exists in the nal (terminal) nodes [33]. The resulting tree can then be pruned to prevent overtting and to obtain the most accurate model for prediction [33], [35]. The third algorithm was a GBM, which creates multiple weakly associated decision trees that are combined to provide the nal prediction [35]. This technique, known as `boosting' can often give a more accurate prediction than a single model [35].

These algorithms were chosen to allow comparison of different commonly used techniques for predictive modelling, with the three specic algorithms being selected to allow comparison of a regression technique (logistic regression), a single decision tree (RPART), and a tree based ensembl technique (GBM). The choice of the three algorithms also allows us to compare the performance of two novel to the area machine algorithms (RPART and GBM) with the more traditional logistic regression model. The three algorithms vary in terms of how the modelling is carried out and the complexity of the nal models. The possibility of practical implementation of the solution was also considered. Characteristics of the dataset were also important in the choice of model. For example, different algorithms are typically used depending on whether the problem is regression or classication, and in this case algorithms suitable for classication were used. The model parameters associated with each algorithm were tuned using ten fold cross validation repeated ve times, over a custom tuning grid. This process identies the optimal tuning parameters, and helps to prevent against overtting. For logistic regression there are no tuning parameters, but resampling was still performed to evaluate the performance of the model [35]. The tuning parameters commonly used for 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 [35]. 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 [35]. To further prevent against overtting 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,

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

sensitivity and specicity. 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 [36].

IV. RESULTS

A. DESCRIPTIVE STATISTICS

Table 1 presents the descriptive statistics for the dataset. Across both hospitals, 24% of the ED attendances resulted in an admission to hospital, with 26.5% of attendances resulting in an admission at hospital 1 and 19.81% at hospital 2. This compares similarly to other hospitals in Northern Ireland and England [37], [38]. Similar admission rates can also be observed at hospitals internationally with studies carried out in Singapore where 30.2% of ED attenders were admitted [8], in Canada where 17.9% of ED attenders were admitted [22] and in the USA where 34% were admitted [25]. However, some of these studies relied on single hospital sites or a small number of hospitals, which could be unrepresentative of national admission rates. Whilst the admission date was disaggregated into the day, week, and month, the week of the year was not included in the nal models as it reduced the performance of the model. Overall, attendances and admissions were higher on weekdays than at weekends with the highest number of admissions being on Mondays. Baker [14] observes a similar trend in England, with the highest frequency of attendances on Mondays and decreasing attendances through to Friday. However, Baker [14] also shows that attendances slightly increased at the weekend with Sunday being the second busiest day. ED attendances are lowest in the winter months and.

TABLE 2. Model performance.

	Accuracy (%)	Kappa	AUC-ROC	Specificity	Sensitivity
Logistic Regression	79.94	0.4600	0.8497	0.8995	0.5357
Decision Tree (RPART)	80.06	0.4661	0.8249	0.9015	0.5349
GBM	80.31	0.4724	0.859	0.9038	0.5379

This study used a data mining approach to develop and assess three machine learning algorithms to predict the robability of admission at the point of triage. Overall, the results show that the GBM performed best, although the decision tree and logistic regression models only performed slightly less well, thus making all three models potential candidates for implementation. Although the GBM was the most accurate of the three models, in scenarios where interpretability is important logistic regression model may be the most promising candidate for implementation due to its simplicity and ease of interpretation. This follows the process recommended by Kuhn and Johnson [35]. They propose three steps for identifying an implementable model: 1. Build the potentially most accurate model using complex and less interpretable models; 2. Build simpler models using more interpretable algorithms; 3. If the accuracy of the simpler model is suf- cient compared to the more complex model consider this model for implementation. In this study, the simpler models (logistic regression and the decision tree) compare quite well with the more complex GBM. The logistic regression model is also straightforward to interpret and understand and learly articulates howdifferent factors are contributing to the

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

prediction, which may assist with clinician buy in and condence in the prediction. Whilst decision trees can be interpreted, they can be unstable with small changes in the data potentially drastically changing the structure of the tree [41]. Ensembles of decision trees, such as GBM's, can be similarly difcult to interpret as they combine multiple single decision trees to derive the nal predictions. However, in scenarios where accuracy is paramount, theGBMwould be the optimal choice for implementation. The models presented in this study have higher levels of accuracy when compared to several other studies presented in the literature. For example, using logistic regression to model data held on the hospital administrative systems about patients aged over 75, LaMantia et al. [9] achieved an AUC-ROC of 0.73. They postulate that their model is not accurate enough by itself to make an individual level admission decision. Using logistic regression, Sun et al. [8] achieved similar accuracy to the models presented here, w ith an AUC-ROC of 0.849. It is notable that Sun et al. [8] do not achieve higher accuracy than the models presented here despite including data about pre-existing conditions. They found that admission was more likely for patients with diabetes, hypertension and dyslipidaemia.

However, Cameron et al. [11] achieved a slightly higher accuracy using a logistic regression model, with an AUC-ROC of 0.8774. They included two variables which were unavailable in this study: the national early warning score (NEWS), which is not used in Northern Ireland; and the referral source, which isn't always captured at the ointriage in Northern Ireland. They also covered a larger geographical area, and consequently had a larger sample, which could also have improved the accuracy of their model. The analysis of the descriptive statistics and logistic egression model also highlights some important patterns in data. Admissions are linked to the patient's age, arrival mode, triage category, care group, previous admissions, the hospital and to a lesser extent temporal variables. Although the results show that admission is more likely with more severe triage categories, the descriptive statistics also highlight the potential for admission across the categories. Potential explanations for this could be that patients deteriorate after being triaged, or that additional information relating to their condition becomes available, resulting in an admission.

The logistic regression model also highlights that admission is more likely when patients arrive by ambulance. This may be due to the increased propensity for patients to call an ambulance for more serious conditions. This compares similarly to other studies which have also identied a positive relationship between arrival by ambulance and admission to hospital [8], [11]. Similarly, the care group and triage category are likely to be proxies for the severity of the patient's condition. It is also possible that patients with different types of conditions attend different ED's at different times, which could account for the signicance of temporal and site differences. Although these relationships are interesting and useful in informing the model development process, the overall aim of the study was not to gain inference, but to develop predictive models. Further research would therefore be required to conrm any underlying causal mechanisms.

Vii 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 esources based on the expected patient in_ow from the ED. This could help to improve patient _ow 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 omparing 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 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.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.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. VOLUME 6, 2018 10467**
B. Graham *et al.*: Using Data Mining to Predict Hospital Admissions From ED
[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/brie_ng-papers/sn06964.pdf
[15] DHSSPS. (2015). *Northern Ireland Hospital Statistics: Emer-gency Care (2014/15), Northern Ireland*. [Online]. Available:
https://www.dhsspsni.gov.uk/sites/default/_les/publications/dhssps/hsemergency-care-2014-15.pdf