Studies in Engineering and Technology
Vol. 10, No. 1; 2023
ISSN 2330-2038E-ISSN 2330-2046
Published by Redfame Publishing
URL: http://set.redfame.com

A Machine Learning Framework for Length of Stay Minimization in Healthcare Emergency Department

Daniel Asuquo*[1,2], Imeh Umoren[3], Francis Osang[4] & Kingsley Attai[5]

[1]Department of Computer Science, Faculty of Science, University of Uyo, PMB 1017, Uyo, Nigeria

[2]TETFund Centre of Excellence in Computational Intelligence Research, University of Uyo, Nigeria

[3]Department of Computer Science, Akwa Ibom State University, Ikot Akpaden, Mkpat Enin, Nigeria

[4]Department of Computer Science, National Open University of Nigeria, Abuja, Nigeria

[5]Department of Mathematics and Computer Science, Ritman University, Ikot Ekpene, Nigeria

Correspondence: Daniel Asuquo, Department of Computer Science, Faculty of Science, University of Uyo, PMB 1017, 520003, Uyo, Nigeria. E-mail: danielasuquo@uniuyo.edu.ng

Received: November 22, 2021 Accepted: December 13, 2021 Online Published: September 8, 2023

doi:10.11114/set.v10i1.6372 URL:https://doi.org/10.11114/set.v10i1.6372

Abstract

The emergency departments (EDs) in most hospitals, especially in middle-and-low-income countries, need techniques for minimizing the waiting time of patients. The application and utilization of appropriate methods can enhance the number of patients treated, improve patients' satisfaction, reduce healthcare costs, and lower morbidity and mortality rates which are often associated with poor healthcare facilities, overcrowding, and low availability of healthcare professionals. Modeling the length of stay (LOS) of patients in healthcare systems is a challenge that must be addressed for sound decision-making regarding capacity planning and resource allocation. This paper presents a machine learning (ML) framework for predicting a patient's LOS within the ED. A study of the services in the ED of a tertiary healthcare facility in Uyo, Nigeria was conducted to gain insights into its operational procedures and evaluate the impact of certain parameters on LOS. Then, a computer simulation of the system was performed in R programming language using data obtained from records in the hospital. Finally, the performance of four ML classifiers involved in patients' LOS prediction: Classification and Regression Tree (CART), Random Forest (RF), K-Nearest Neighbour (K-NN), and Support Vector Machine (SVM), were evaluated and results indicate that SVM outperforms others with the highest coefficient of determination (R^2) score of 0.986984 and least mean square error (MSE) value of 0.358594. The result demonstrates the capability of ML techniques to effectively assess the performance of healthcare systems and accurately predict patients' LOS to mitigate the low physician-patient ratio and improve throughput.

Keywords: Length-of-stay prediction, healthcare, emergency department, overcrowding, resource allocation, machine learning

1. Introduction

The emergency department (ED) is normally the initial point where healthcare services are accessed for many hospital visits, and in that process becomes the starting point of unplanned episodes of patient care. According to Rashwan, Habib, Arisha, Courtney & Kennelly (2016), the key functions of the ED include the provision of immediate attention to the acutely unstable patients and timely treatment of patients with non-life-threatening cases so patients do not remain more than 12 hrs. The ED is regarded as a complex service system where most patients arrive with little or no information about their health conditions or requirements. Therefore, the caregivers are expected to assess and classify patients according to their emergency severity index or acuity level to facilitate treatment or disposition (Oddoye, Jones, Tamiz, & Schmidt, 2009). In most countries, the review and treatment of each patient, and their appropriate disposition must occur within service time targets established by government and regulatory agencies (Geue et al., 2012). However, the random nature of patients' arrival coupled with overcrowding due to inadequate healthcare facilities and poor physician-patient ratio pose serious challenges to ED timely service delivery and overall system performance. Forero et al. (2010) reported that overcrowding in ED is the major cause of delays in care which in turn increases patients' waiting time and length of stay (LOS). It is also noted to negatively affect patients' safety causing discomfort and dissatisfied experience in the hospital (Abo-Hamad and Arisha, 2013; Graff, 1999). The growing demand for EDs despite budget and cost constraints has made healthcare decision-makers to be on continuous pressure to optimize, control, and manage their resources in a more effective and efficient way. Specifically, doctor-to-patient ratio as well as nurse-to-patient ratio is low in most sub-Saharan countries, and World Health Organization (WHO) data reveals doctors per 1,000 persons as: Nigeria - 0.3, Uganda - 0.12, and South Africa - 4.3. Physician burnout is certain with a detrimental impact on patients' LOS, physician safety, and overall healthcare performance. Burnouts as high as 75.5% are reported among physicians in Nigeria (Nwosu, 2020). The WHO expresses concerns about the scarcity of resources in these countries impeding the attainment of Sustainable Development Goals 3, Target 3.3.

According to Bagust, Place, and Posnett (1999), Thorwarth, Rashwan, and Arisha (2015), and Yousefi, Yousefi, Ferreira, Kim, and Fogliatto (2018), optimal resource planning is required for all ED admissions to facilitate the reduction in staff burnouts which are responsible for avoidable revisit and left without being seen (LWBS) situations. This will help maintain an acceptable workload for nurses and prevent most of them from yearning for other departments (ACEP, 2010). Nevertheless, sudden changes to the workload due to the complex dynamic nature of the ED system caused by emergencies such as fire, flood, natural disasters, disease outbreaks (e.g. Ebola, COVID-19), and terrorist attacks are difficult to predict. A typical ED system is composed of a collection of resources including humans (e.g. doctors, nurses, and technicians) and equipment (e.g. X-ray machines and CT-Scan). The system involves human processes and decision-making to determine how it evolves based on the awareness of the situation and available resources. Therefore, appropriate steps must be taken to ensure performance and quality improvement. Barnes, Hanson, and Giraud-Carrie (2018) advocate the introduction of computational health science through the interdisciplinary application of innovative computer science tools to address health-related questions and problems. In Shafaf and Malek (2019), a review of the application of machine learning (ML) approaches in emergency medicine was presented. The authors explained that as the number of patients visiting the ED increases, common traditional techniques are no longer sufficient for predicting patient admission, discharge, and triage.

ML computation has emerged as an effective alternative for handling imbalanced data, high-dimension noise reduction, and new complexity that may arise in the ED environment (Mariki, Mkoba, & Mduma, 2022; Hunter-Zinck, Peck, Strout, & Gaehde, 2019) and other problem areas (Attai., Amannejad, Vahdat, Pour, Obot & Uzoka, 2022; Asuquo, Umoh, Osang, & Okokon, 2020). Integrating ML, simulation, and optimization into a predictive analytic decision framework will facilitate the provision of effective, efficient, and quality healthcare to ED patients. This paper presents an ML framework for improving the efficiency of ED services by optimizing the use of resources such as staff, treatment spaces, and equipment for reduced patients' LOS and increased satisfaction. The framework comprises four ML models: Classification and Regression Tree (CART), Random Forest (RF), K-Nearest Neighbour (K-NN) and Support Vector Regression (SVM) whose performance is evaluated to determine the best predictive model for the task of minimizing patients' LOS using Mean Absolute Error (MAE), Mean Square Error (MSE), Root Mean Square Error (RMSE) and Coefficient of Determination (R^2) metrics. The approach is similar to a flexible manufacturing environment where each job (patient) has different processing requirements and different jobs (patients) compete for access to scarce resources (e.g. beds, doctors, CT-Scan).

The rest of the paper is organized as follows. Section two reviews related literature on the dynamics of hospital ED, challenges of delay resulting in long patients' waiting and care time as well as approaches to mitigating the problem. This is followed by a brief description of selected ML models. Section three presents the proposed ML predictive framework for LOS minimization, along with the nature of collected data for model training and testing as well as performance evaluation metrics. Section four discusses results obtained from simulation in R Programming software while Section five concludes the paper.

1.1 Motivation and Principal Contributions

Big data analytics in ED service facilitates the ability to model ED operations and system dynamics. Also, modeling dynamic patient characteristics and treatment patterns can give insight into effective task scheduling and resource utilization. Although several studies have considered the use of intelligent and ML techniques in LOS optimization in hospital EDs, none of these studies consider the comparison of deterministic models - CART, RF, K-NN, and SVM. Also, most of them evaluated the performance of their models with only one metric (Rashwan etal. 2016; Yousefi et al. 2018) while this work compares the prediction accuracy of the four ML models using four metrics – MAE, MSE, RMSE, and R^2. Furthermore, this work presents an ML framework for predicting patients' LOS following implementable treatment pathways for efficient and robust ED resource allocation. The adaptation of big data analytics also facilitates precise categorization that helps determine the correlations, hidden patterns, and other valuable insights from the vast amount of data with varied properties through the classification process.

2. Related Work

The increased advancement in computer and network technologies has made the development of useful information systems for every human endeavor possible. Consequently, health information system has brought great advances in recent years, which result from the use of techniques of artificial intelligence and ML in developing a more complete, more accurate, and more robust healthcare system (Obot, Asuquo, Attai, Johnson, Arnold, Edoho, Ekpenyong, Akwaowo, Udoh, Usen, & Uzoka, 2023; Hashmi, Abidi, & Cheah, 2002; Tomar and Agarwal 2013; Yoo et al 2012; Gül and Guneri, 2015). Magoulas and Prentza (2001) highlighted the importance of ML techniques as data analytics tools for providing more efficient monitoring, detection, and alarming services to doctors and patients. Currently, the trend and concern for hospital top management is to enhance the performance of medical services (Chonde, Parra, & Chang, 2013; Poulymenopoulou, Malamateniou, & Vassilacopoulos, 2008), by maximizing the utilization of scarce medical resources. This optimization is expected to result in minimized patient LOS defined as the total time that a patient spends in the ED from arrival to departure.

The work of Priya, AnandhaKumar, & Maheswari (2008) as well as Yeh and Lin (2007) used metaheuristics to solve the scheduling problem of a hospital to optimize the schedule of doctors, nurses, or patients. The waiting time of patients was reduced as metaheuristics provided a better work plan. The main issue with this approach is that, in addition to input and output, the transition, evaluation, and determination operators have to be performed repeatedly until the search process converges or meets the predefined stopping condition (Tsai &Rodrigues, 2014). Manupati, Teja, Hussain, Sandeep, & Varma (2015) developed a linear programming mathematical model to solve the problem of patient admission scheduling aimed at reducing patient waiting time by improving the use of resources. They later adopted a multi-objective Non-dominant sorting genetic algorithm to optimally solve this problem since the linear model had no closed-form solution. Tsai, Chiang, Ksentini, and Chen (2016) provide a brief survey on metaheuristics and emphasized the essence of big-data analytics framework for healthcare systems. Othman and Hammadi (2017) formulated a fitness function using an evolutionary algorithm and fuzzy logic suitable for building a decision support tool for healthcare task scheduling in Pediatric ED. They were able to predict specific limits for the optimal values of the criteria to solve the problem of peaks of activity and overcrowding as well as improve system performance and patient satisfaction. Umoren, Udonyah, and Isong (2019) proposed a computational intelligence framework to predict patient's LOS in hospital ED. However, they analyzed several factors including Severity of Illness or Emergency Cases (SIC) to assess its performance but the ML framework was implemented using Intuitionistic Type-2 Fuzzy Logic System (IT2-FLS). In Yousefi et al. (2018), a meta-model comprising an ensemble of Adaptive Neuro-Fuzyy Inference System (ANFIS), Feed Forward Neural Network (FFNN), and Recurrent Neural Network (RNN) was used for optimum resource planning in the ED. The results were compared and evaluated in terms of mean absolute percentage error (MAPE) only. In Taboada, Cabrera, Iglesias, Epelde, & Luque (2011), an agent-based decision support system was developed for a hospital's ED. Rashwan et al. (2016) applied a model that integrates three approaches: simulation, multivariate factor analysis, and multi-objective optimization to support management decisions on key parameters affecting the treatment journey of patients. A set of supervised ML predictive models was tested to select the best surrogate model for each response variable. Only root mean squared error (RMSE) was used to evaluate the effectiveness of the regression models.

From the foregoing, improving the performance of EDs is vital to the success of the healthcare organization. Moss and Xiao (2004) attempted to improve the ED system by capturing ED workflow patterns and analyzing these patterns to create an automated and enhanced ED system design. Hashimoto and Bell (2007) developed simulation tools to assist healthcare decision-makers in this endeavour. Another approach was the use of workflow technologies and web services to automate emergency healthcare processes (Poulymenopoulou et al., 2008). While some attempts have been made to handle issues about resource allocation, task scheduling, service cost, and length of stay minimization, research is still needed in the use of ML models to predict the perceived output (e.g. LOS) for a given set of input predictors. This paper compares the performance of four supervised ML techniques in the task of prediction of patients' LOS. The outcome has the potential implication to guide the hospital's effort to optimize its ED services towards individualized care delivery thereby maximizing throughput, reducing waiting time, and abetting LWBS situations.

2.1 ED Key Performance Indicators

In most cases, the human resources in the ED include a receptionist, doctors, nurses, and technicians. Equipment available may include bed space, oxygen machine, X-ray machine, CT scanner, laboratory, and pharmacy. The ultimate goal of the ED is patient satisfaction, which is normally evaluated by its LOS. The total time spent in the ED from arrival to departure is referred to as the patient's LOS. This is divided into two key phases: delay time which is prior to the commencement of treatment, and care time which runs from the start of treatment until the eventual departure from the ED. These phases are marked in Figure 1 along with the time stamps present in the ED profile that determines the different stages. Performance targets for the delay time and the ED LOS are often defined by government and regulatory agencies. Oftentimes, triage is used to determine the priority of ED patients based on the severity and urgency of their condition.

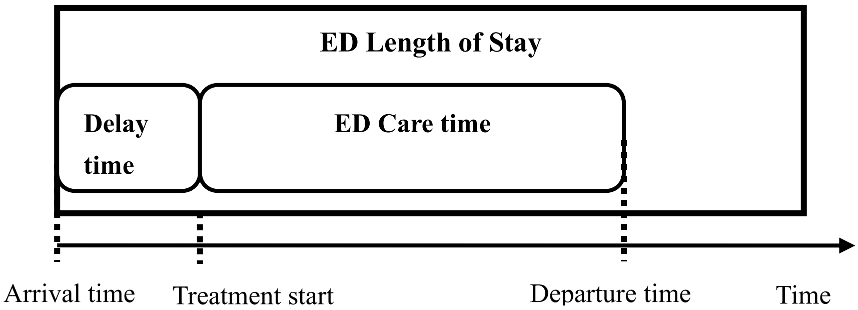

Figure 1. LOS components: delay time and care time

Certain care or service quality metrics must be collected and analysed at the ED to determine areas of improvement necessary to meet set goals. These metrics are categorized into two major key performance indicators (KPIs) - time measures and proportion measures (Welch, Augustine, Camargo, & Reese, 2006; Ghanes, Diakogiannis, Jouini, Jemai, & Wargon, 2014). Time measures include average waiting time, average care time, average LOS, doctor to decision to admit time, arrival time to rooming, disposition to discharge, etc. The proportion measures include a number of LWBS cases, number of discharged patients, number of those who left before they were supposed to, complaints, hospital diversion, and ED patient flow, etc. Data on these core measures have to be collected, stored, and evaluated using advanced analytical tools where comparative reports with regard to national averages can be generated. Such results help hospital managers to proactively assess performance and identify opportunities for quality improvement. However, certain demography-dependent issues have caused most hospital EDs to witness unsatisfactory levels of these service quality metrics. For instance, inadequate patient beds always result in long arrival to discharge times, which increases patient's LOS. The lack of a CT-Scan machine and nurse shortage may also impact negatively on overall ED performance.

2.2 Description of Selected ML Models

(a) Classification and Regression Tree

The classification and regression tree (CART) is a largely used non-parametric ML technique for effectively solving regression and classification problems. It is a decision tree that uses if-then-rules to get solutions by making sequential, hierarchical decisions about the outcome variable based on the predictor data. CART builds classification models in the form of a tree structure by breaking down a dataset into smaller and smaller subsets while at the same time developing an associated decision tree incrementally. The final result is a tree with decision nodes and leaf nodes (Saxena, 2017)**.**

In the tree structure, an internal node represents a feature (attribute) while the branch represents a decision rule, and each leaf node represents the outcome useful for decision-making. The topmost node - the root node, learns to recursively partition the tree based on the attribute value. The attributes are chosen for the root node and other sub-nodes while keeping the highest information gain and low Gini index. The time complexity of decision trees is a function of the number of records and number of attributes in the given data. Decision trees can handle high-dimensional data with good accuracy. It is easy to visualize, understand, and interpret the results. The CART structure, shown in Figure 2, is based on the following steps:

STEP 1: Select the best attribute using Attribute Selection Measures (ASM) such as Gini Index to split the records;

STEP 2: Make that attribute a decision node and break the dataset into smaller subsets;

STEP 3: Start tree building by repeating this process recursively for each child until one of these conditions matches:

(i) All the tuples belong to the same attribute value

(ii) There are no more remaining attributes

(iii) There are no more instances

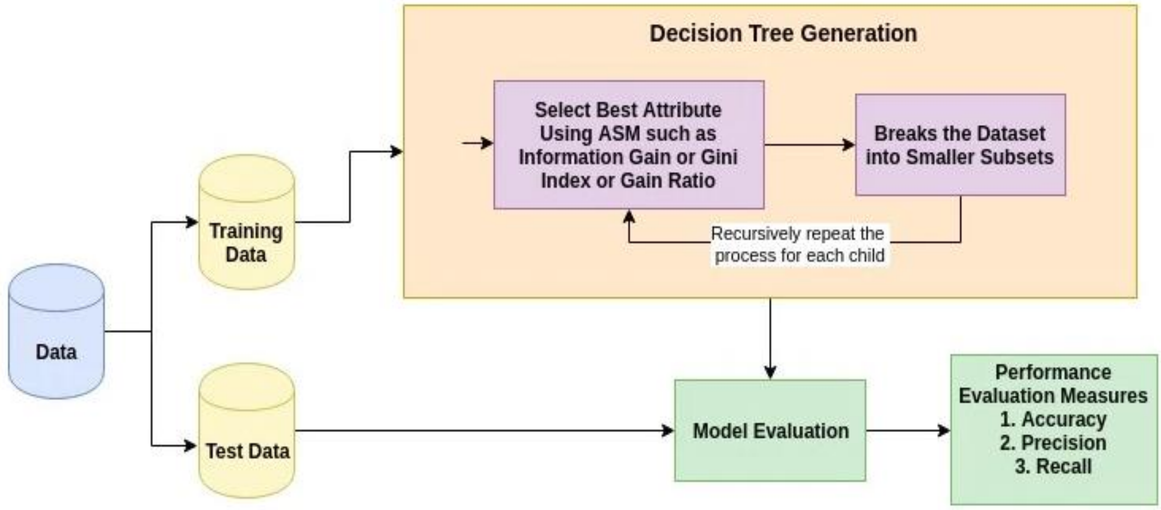

Figure 2. Simplified CART structure for decision tree generation

(b) Random Forest

Random forest (RF) is an ensemble ML algorithm used for classification and regression problems. It applies the technique of bagging (bootstrap aggregating) to decision tree learners. Thus, as an ensemble learning algorithm, the final prediction, in classification, is the average of the most frequent prediction. So, the algorithm takes the average of many decision trees to arrive at a final prediction. Furthermore, it is a relatively easy model to build and does not require much hyperparameter tuning.

The essential features of RF algorithm include:

1. Ensemble learning prevents over-fitting of data
2. Bootstrapping enables random forest to work well on relatively small datasets
3. Predictors can be trained in parallel
4. Decision tree learning enables automatic feature selection

Bagging is simply a method of generating new datasets from existing data by creating samples of the existing data with replacement. This means there could be repeated values in each of the newly created datasets. This process helps RF avoid overfitting, despite increasing the number of trees. This is because it averages many low-bias and high-variance predictors, thereby reducing the variance without increasing bias. Also, since multiple versions of the dataset are generated, it is possible to work with relatively small datasets. The pseudocode for classification tasks with RF is presented in Figure 3.

STEP 1: Randomly select ***k*** features from a total of ***m*** features, where $k<<m$

STEP 2: Among the ***"k"*** features, calculate the node ***"d"*** using the best split point

STEP 3: Split the node into **daughter nodes** using the **best split**

STEP 4: Repeat steps **1 to 3** until "l" number of nodes has been reached

STEP 5: Build a forest by repeating steps ***1 to 4*** for "n" number of times to create ***"n" number of trees***

Figure 3. Pseudocode for classification with RF algorithm

(c) K-Nearest Neighbour

The K-Nearest Neighbour (K-NN) is a simple and easy supervised ML algorithm. As a non-parametric method, it assumes that similar inputs have similar outputs and the method can be used for both regression and classification. It has a minimal training phase as most of the effort is expended in the testing phase. For that purpose, it is said to be an instance-based as well as a lazy learning algorithm since it attempts to obtain the best prediction at every instance of the testing phase. Its interpretation is easy and requires less calculation time than other ML algorithms like CART and RF. In K-NN classification, new samples are classified by assigning the class that is the most common among the *k* closest sample in the training set. To determine the closest sample, some form of distance measure such as Euclidean, Manhattan, Minkowski, or Hamming distance is required.

(d) Support Vector Machine

SVM can be used for both classification and regression tasks (Smola & Sch¨olkopf, 2004). As a binary classifier, it divides its input space into two regions, separated by a linear boundary. Ideally, SVM, unlike other classifiers, increases the confidence of classification by maximizing the decision surface. This allows a clear separation of data in the data space and easy selection of data precisely to one class or the other (Wang &Lin, 2014) using a margin or hyperplane. SVM is a very useful ML tool for learning linear predictors in high-dimensional feature spaces with the capacity to handle both sample complexity and computational complexity challenges. The algorithm deals with the sample complexity challenge by searching for "large margin" separators. It facilitates the mapping of data in the feature space with very less computation in comparison to other classification algorithms.

For high dimensional space, kernel functions are used for defining such hyperplanes to separate classes nonlinearly. The kernel function transforms the input data into the desired output form. It separates linearly non-separable data into the linearly separable data. Kernel based learning algorithms, and in particular kernel-SVM, are very useful ML tools and their success may be attributed both to being flexible for accommodating domain specific prior knowledge and to having a well-developed set of efficient implementation algorithms. Although different SVM kernel functions abound, the Gaussian Radial Basis Function happens to be the most widely applied. The Support Vector Regression (SVR) algorithm used to predict the output, patient's LOS, through a non-linear SVR model is presented in Figure 4 and described as follows (Smola, O´va´ri, & Williamson, 2001; Sch¨olkopf & Smola, 2002):

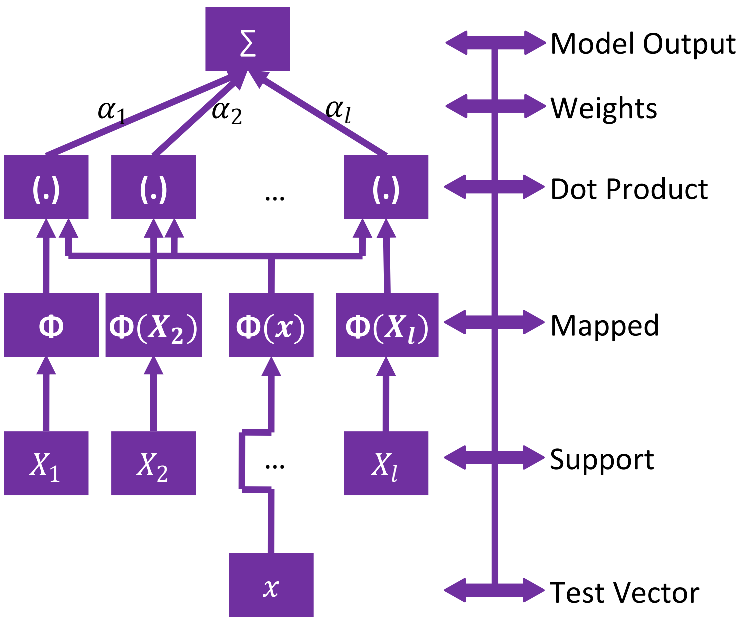

Figure 4. Structure of SVR model

Given a dataset with n dimensional features and a target variable $\{(X_1, y_1),\ (X_2, y_2), \dots, (X_m, y_m); i = 1, \dots, m\}$, where $X \epsilon R^n$, $y \epsilon R$. The objective is to find a function $f(x)$ with at most ε-deviation from the observed target y. Since the relationship between X and y is non-linear, a non-linear SVR model formulated as a maximization problem given as follows:

$$\max\left\{\frac{1}{2}\sum_{i=1,j=1}^{m}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)\langle\Phi(X_i), \Phi(X_j)\rangle - \varepsilon\sum_{i}^{m}(\alpha_i + \alpha_i^*) + \sum_{i}^{m} y_i\,(\alpha_i - \alpha_i^*)\right\}$$

Such that:

$$\textstyle\sum_{i=1}^{m}(\alpha_i + \alpha_i^*) = 0; 0 \leq \alpha_i, \alpha_i^* \leq C \quad (1)$$

where, α_i and α_i^* are the model weights, ε is epsilon, and C is the complexity and number of support vectors.

The dot product is computed in Eq. (2) as:

$$\big(\phi(x).\phi(X_i)\big) = K(x, X_i) \quad (2)$$

where, $\phi(X_i)$ and $\phi(x)$ are the mapped vectors.

The $\phi(X_i)$and $\phi(X_j)$ mapping functions are computed using radial basis function kernel, $K(x, X_i)$ using eq. (3) as follows:

$$K(x, y) = \exp\left(-\frac{1}{2\sigma^2}||x - y||^2\right) \tag{3}$$

The output of the SVR algorithm, which is the predicted patients' LOS, is obtained as expressed in Eq. (4).

$$y_i = \sum \alpha_i K(x, X_i) + b \tag{4}$$

where, y_i is the predicted patient's LOS, α_i is the model's weight; b is the bias; and $K(x, X_i)$is the kernel function.

3. Methodology

3.1 ML Framework for Patients' LOS Prediction in ED

To aid decision-makers in the ED in allocating available resources effectively and efficiently to ensure quality healthcare service delivery, prevent overcrowding and reduce waiting time subject to capacity as well as budget constraints, the performance of four supervised ML models (CART, RF, K-NN, and SVM) is compared to determine which is best in the task of predicting patients' LOS. Figure 5 depicts the framework for optimizing patients' LOS in the hospital ED where the formulated optimization problem was solved through a discrete-event simulation using R programming language. A raw data of about 1146 patients was used with needed features extracted and split into training and testing dataset in the ratio of 7:3. All these form part of the big-data analytics module.

The predictive analytics decision module handles the simulation of the models and subsequent transformation to the various learned models. The predicted results are visualized and the performance of the classifiers is evaluated to determine which outperforms others based on selected metrics. The predicted results can be presented graphically for visualization and interpretation in the form of histogram of residuals, normal Quantile-Quantile (Q-Q) plots, scatter plots and bar charts for proper comparison of predicted and tuned values. Metrics used for performance evaluation are the MAE, MSE, RMSE and R^2. From the framework, prioritized actions can be taken and useful recommendations made to aid decision-making and facilitate refinement of the processes geared towards efficient resource allocation in the hospital's ED, taking into consideration typical treatment pathways, shown in figures 6 - 7. The mathematical expressions for MAE, MSE, RMSE, and R^2 are as given equations (5) - (8).

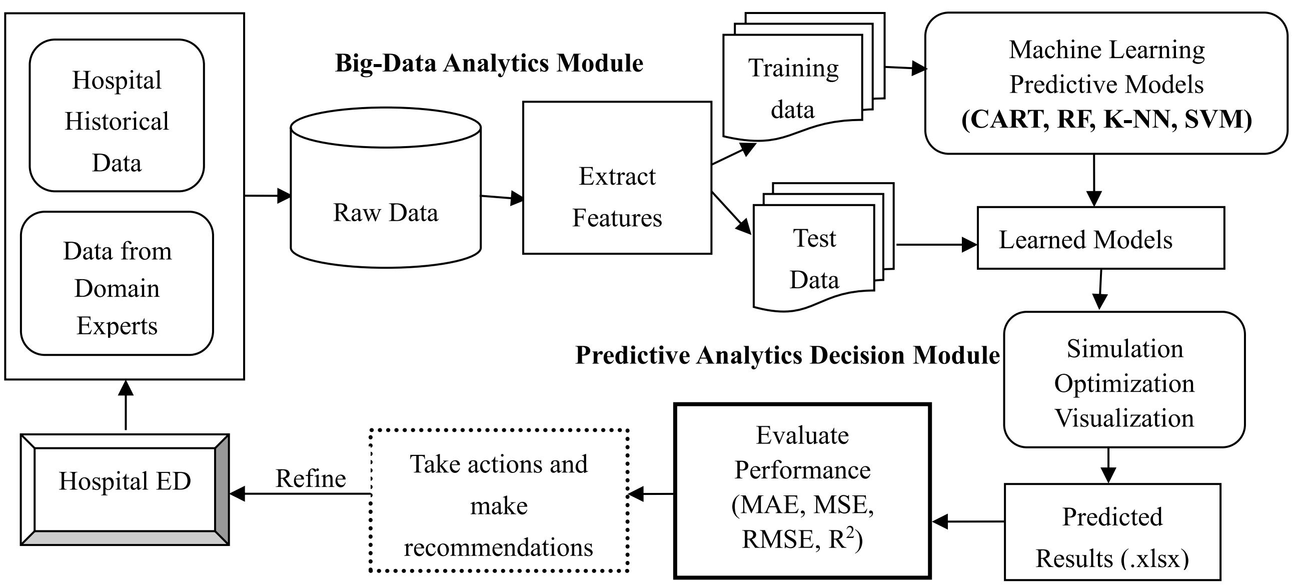

Figure 5. Proposed ML framework for minimizing patients' LOS in hospital ED

$$MAE = \frac{|y_1 - \widehat{y_1}| + |y_2 - \widehat{y_2}| + \ldots + |y_m - \widehat{y_m}|}{n} \tag{5}$$

$$MSE = \frac{(y_1 - \widehat{y_1})^2 + (y_1 - \widehat{y_1})^2 + \ldots + (y_m - \widehat{y_m})^2}{n} \tag{6}$$

$$RMSE = \sqrt{MSE} \tag{7}$$

$$R^2 = \left(\frac{n\sum xy - \sum x \sum y}{\sqrt{n\sum x^2 - (\sum x)^2} * \sqrt{n\sum y^2 - (\sum y)^2}} \right)^2 \tag{8}$$

where, y_i are the actual observed values and $\widehat{y_i}$ are the predicted values; x is the independent variable while y is the dependent variable. As a measure of goodness of fit of a model, R^2 gives the proportion of the variance in the dependent variable that is predictable from the independent variables. The higher the R^2 score and the lower the error measures, the better the prediction model.

Residuals are estimates of experimental error derived from the difference between observed and predicted responses. The predicted response is obtained from the chosen models, after estimating the model parameters from given experimental data. Examining residuals is a key part of all statistical modeling and inference. The residuals should be approximately normal and independently distributed with zero mean and constant variance.

3.2 Treatment Pathways for Efficient ED Resource Allocation

This section presents two flow diagrams shown in figures 6 and 7, describing two typical treatment pathways that can be used in ED resource allocation system. For clarity, the nodes are numbered and their corresponding task description, resource and duration are listed in Tables 1 and 2. From these scenarios, the least possible patient's ED care time which equals the patient's LOS in the ED is an hour plus 50 minutes, whereas the upper boundary increases (considering ED waiting time) as factors internal and external to the system fluctuates due to dynamic nature of the ED environment. The patient arrival rate to the ED is Poisson distributed while service rate is exponentially distributed. Patients could arrive by self, from general hospitals, private hospitals, home health care facility, and others like ambulance and police custody arrivals.

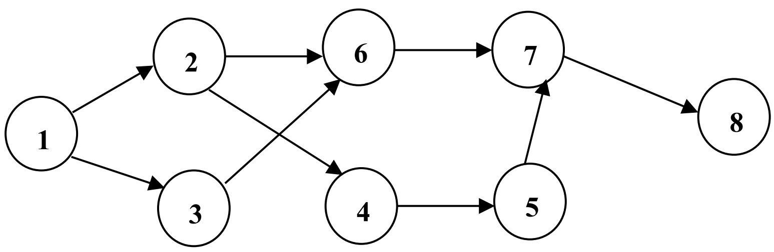

Figure 6. Treatment pathway 1: Medical pathway with admission to hospital

Table 1. Resource and task duration required for treatment pathway 1

Task	Description	Resource	Duration (mins)
1	ED bed allocation	Nurse	15
2	Medical assessment	Doctor	25
3	Vital signs and ECG	Nurse	20
4	Take blood	Nurse	15
5	Pathology	Pathologist	60
6	Writ up patient notes	Doctor	25
7	Admit to inpatient ward	Ward Doctor	15
8	Transfer to inpatient ward	Ward Staff	35

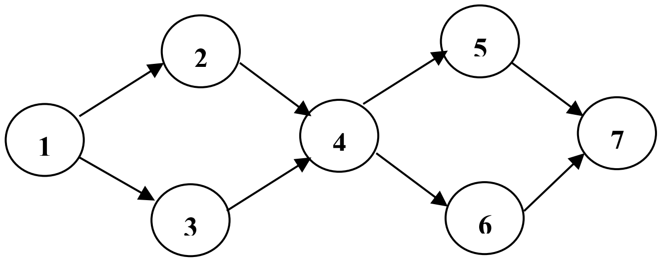

Figure 7. Treatment pathway 2: General medical complaint with discharge home

Table 2. Resource and task duration required for treatment pathway 2

Task	Description	Resource	Duration (mins)
1	ED bed allocation	Nurse	15
2	Medical assessment	Doctor	25
3	Obs+/- Cannulate	Nurse	20
4	Take blood	Nurse	15
5	Pathology	Pathologist	60
6	Treatment	Nurse	35
7	Review and discharge home	Doctor	25

3.3 Data Collection

A dataset comprising 1146 patient records was obtained and used for model prediction. The data was split into two categories for model training and testing as follows (#train, #test) = (802, 342). Important features were extracted from the raw data and subsequently selected for training by each of the predictive models to generate their respective learned models. For instance, a patient's age, level of income and emergency severity index (triage score) were not considered important contributory factors to LOS and therefore were not used in this work. Eight predictor variables and one output variable were considered. The features or predicting factors are number of doctors (NDrt), nurses (NNur), technicians (NTch), X-ray equipment (NXry), CT-Scan machine (NScn), bed space (NBsp), pathologist (NPat), and rate of patient arrival per hour (PArt). The output variable is LOS. Table 3 presents the sample dataset while Figure 8 indicates the level of variable importance in the task of LOS prediction.

Table 3. Sample dataset

PatID	NDrt	NNur	NTch	NXry	NScn	NBsp	NPat	PArt	LOS
4	10	8	2	2	2	6	7	2	11.58824
6	3	6	4	2	0	10	10	1	19.35294
7	5	3	5	2	0	1	8	1	14.41176
9	7	13	3	2	2	2	5	1	8.764706
14	1	2	1	1	1	2	8	3	20.05882
16	4	20	3	1	2	7	7	1	8.764706
17	9	4	2	1	2	2	10	1	11.58824
18	9	8	4	0	2	3	9	1	9.470588
20	2	9	4	0	0	3	7	2	15.82353
22	9	4	5	2	2	5	8	3	10.88235
23	6	4	4	1	0	3	10	1	14.41176
24	8	1	4	1	1	9	9	1	19.35294
28	8	3	3	0	0	9	8	2	20.05882
29	1	6	5	2	1	1	6	1	15.82353
31	4	19	2	0	0	6	8	1	10.88235
32	8	10	4	1	0	1	7	1	9.470588
33	2	5	2	1	2	10	7	3	22.17647
34	3	6	4	0	2	6	10	3	15.11765
38	10	9	3	1	0	9	9	2	13
39	7	5	2	2	2	7	9	3	14.41176

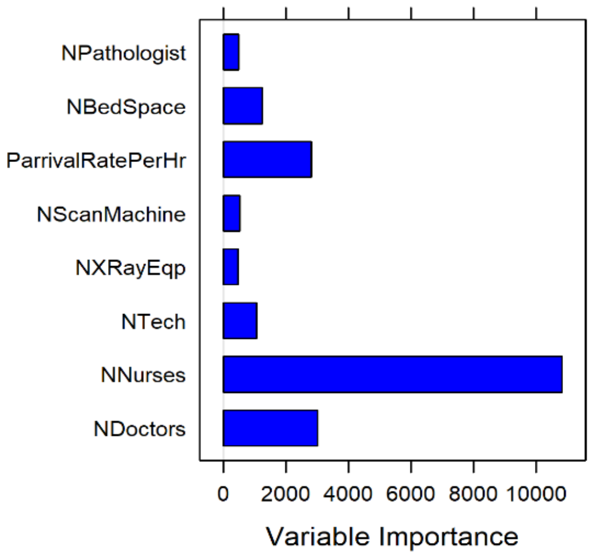

Figure 8. Degree of variable importance

4. Discussion of Results

The simulation model is integrated with statistical analysis component to help identify and visualize the significant factors that affect patient-related performance measures and then optimize them to improve the ED healthcare system. The results from our study include the following for each ML model: actual vs. predicted values, error values, and R^2 score, as well as a histogram of residuals and normal Q-Q plot for the SVM regression model. Figure 9-12 show the actual vs. predicted values for CART, RF, K-NN, and SVM. The result in Figure 9 shows that CART has a poor predictive capability, and high classification error as most of the predicted values do not match the actual values. However, Figures 10 and 11 indicate that RF and K-NN have similar performance in the prediction task while Figure 12 shows that the majority of the actual values were predicted by the SVM regression model.

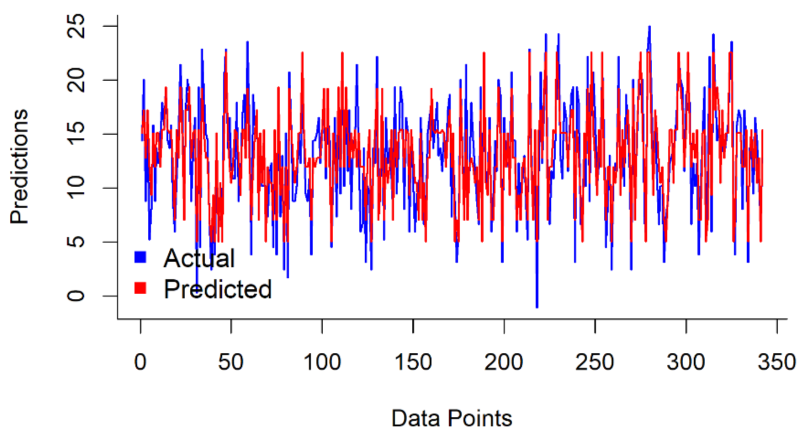

Figure 9. CART Actual vs. predicted result

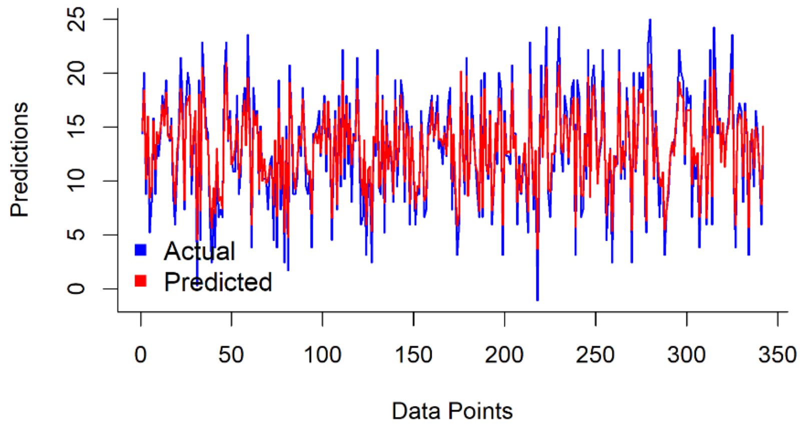

Figure 10. RF Actual vs. predicted result

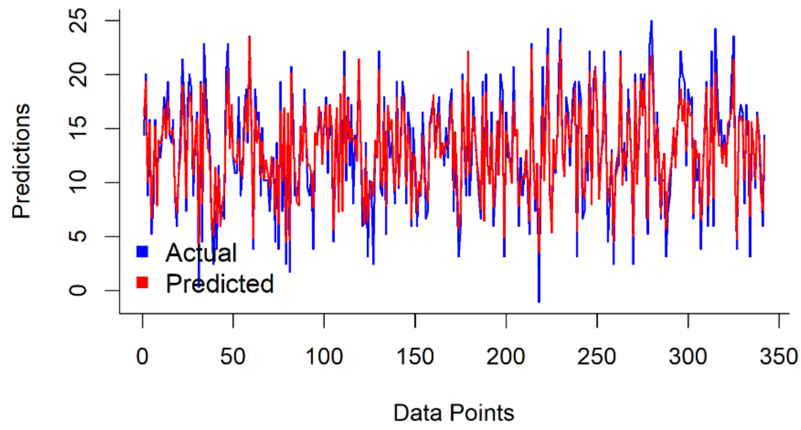

Figure 11. K-NN Actual vs. predicted result

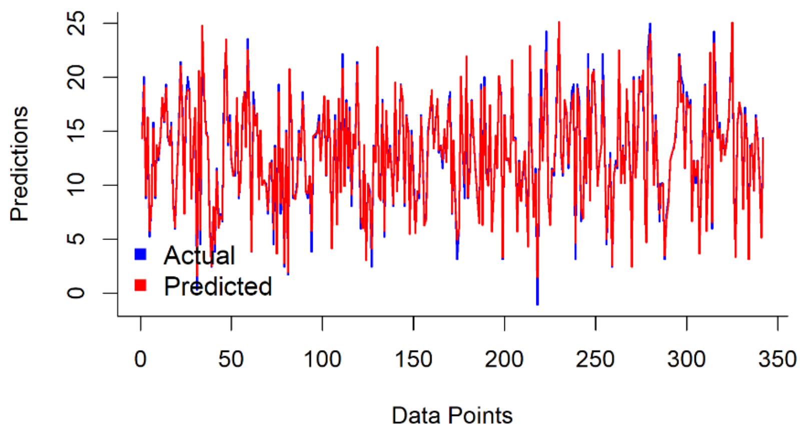

Figure 12. SVR Actual vs. predicted result

The performance of the ML models was evaluated using MAE, MSE, RMSE, and R^2 to determine their predictive accuracy. The error plots for CART, RF, K-NN, and SVM are shown in Figures 13 - 16. In all, results indicate that MAE gave the least error value for each ML model, apart from SVR where MSE was slightly lower. Also, Figure 16 shows that the highest R^2 score was obtained from the SVR model.

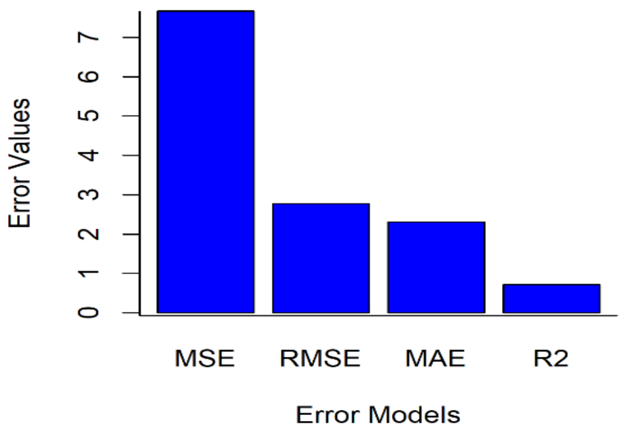

Figure 13. CART error plot

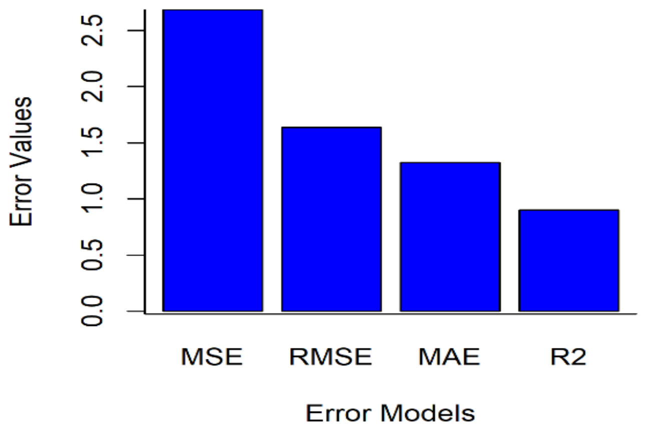

Figure 14. RF error plot

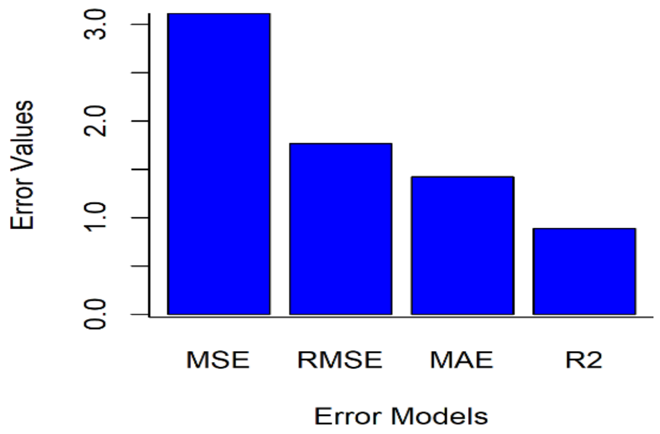

Figure 15. K-NN error plot

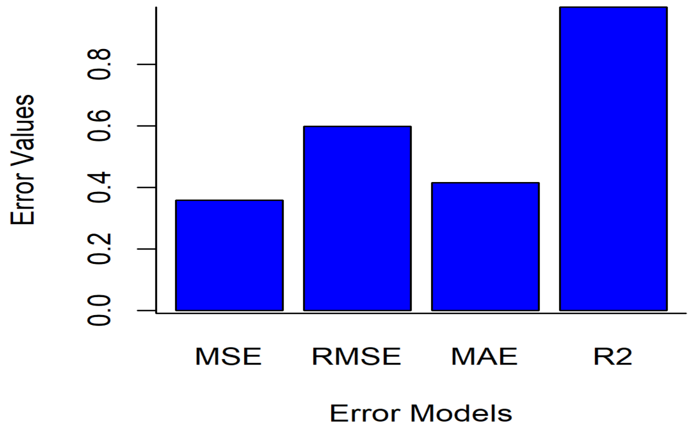

Figure 16. SVR error plot

Each classifier has its strengths and weaknesses; it is assertive to say which one is the best. A comparative analysis of the performance of the ML models in terms of MAE, MSE, RMSE, and R^2 is presented in Table 4. The MAE and R^2 score plots shown in Figures 17 and 18 indicate that SVR outperforms other predictive models in the task of predicting patients' LOS in ED having obtained the least MSE value (0.358594) and highest R^2 score (0.986984). This was closely followed by RF and K-NN while CART had the worst performance.

Table 4. Performance Evaluation of the ML Models

Model/Error Metric	MSE	RMSE	MAE	R^2
CART	7.663155	2.76824	2.300688	0.721843
RF	2.686047	1.638916	1.323623	0.902502
KNN	3.113562	0.764529	1.421790	0.886984
SVR	0.358594	0.598827	0.415986	0.986984

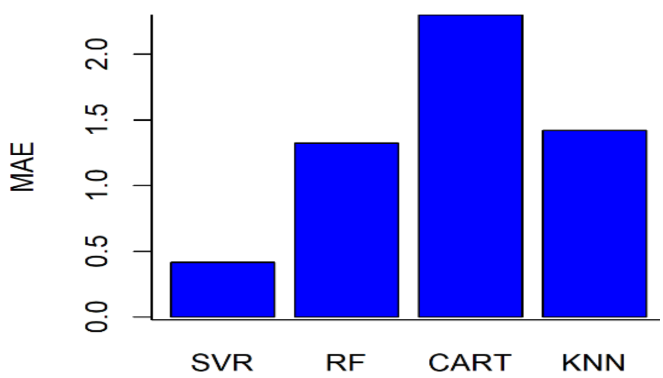

Figure 17. MAE comparative analysis of the ML models

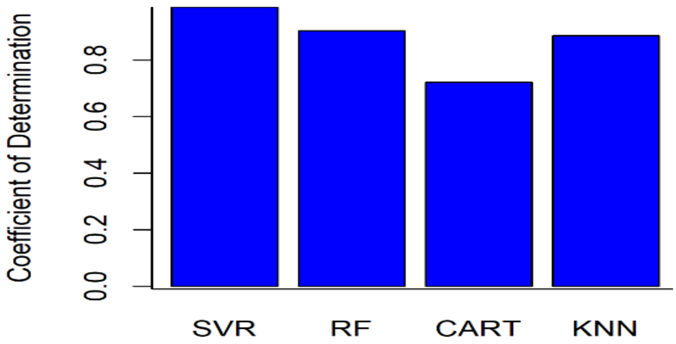

Figure 18. R^2 Comparative analysis of the ML models

The SVR training parameters and values are presented in Table 5 while Figures 19 and 20 show the histogram of residuals and normal Q-Q plots. Ideally, the histogram of residuals is used to check the normality of the data generating process. In Figure 19, the overall pattern of the residuals is similar to a bell-shaped pattern, indicating the regression model is predicting values higher than actual and lower than actual with equal probability. This is acceptable as a typical regression model is expected to err in predicting a response randomly. Consequently, the graph indicates that the errors (residuals of the fitted model) are independent of each other. Furthermore, the Normal Q-Q plot in Figure 20 shows a linear plot of the actual (sample) residual quantiles and the theoretical (perfectly normal distribution) residual quantiles of the same distribution. It simply checks whether the distribution of the residuals is normal or not. If the graph is perfectly overlaying on the diagonal, the residual is said to be normally distributed. In Figure 20, the Q-Q plot looks slightly deviated on the baseline but on both sides of the baseline. This indicates the residuals are distributed approximately in a normal fashion. Superimposed on the plot is a line joining the first and third quantiles of each distribution, where the line is extrapolated out to the ends of the sample to help evaluate the linearity of the data.

Table 5. SVR training parameters and values

Parameter	Description	Value
SVM-Type	eps-regression	-
SVM-Kernel	Radial basis function	-
Cost	C	1
Gamma	γ	0.125
Epsilon	ε	0.1
Number of Support Vectors		237

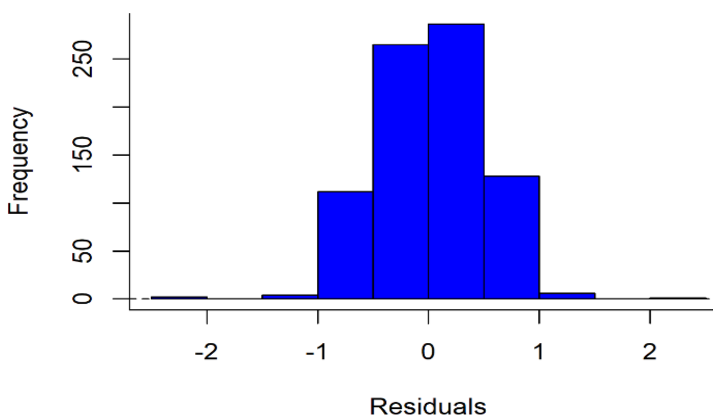

Figure 19. SVR histogram of residuals

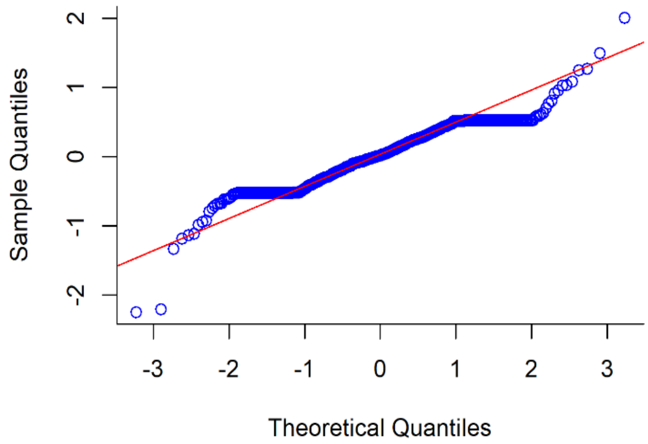

Figure 20. SVR normal Q-Q plot

5. Conclusion and Future Research Directions

The importance of deploying ML techniques to optimize resource allocation, improve service delivery, and minimize the LOS of patients cannot be overemphasized. This has the benefit of mitigating the pressure on ED staff, and reducing overcrowding, amidst inadequate healthcare facilities and low physician-patient ratio. Our proposed framework considered treatment pathways for patients with Poisson arrival rates and exponential service times, and compared the performance of four ML models using MAE, MSE, RMSE, and R^2 score to determine which is best for the task of predicting patients' LOS. Obtained data was split into two for model training and testing, and results from the test data indicate that the prediction model with the most accurate (highest R^2 and least MAE values) performance is SVR, followed by RF, K-NN while CART yielded the worst performance. The result from the ε-SVR predictive model is suitable for deployment in the hospital ED, as it can provide hospital managers a formal assessment of how the ED can cope with the unpredicted increase in workload and demand. Revisits due to inappropriate care coordination following discharge can be avoided and quality healthcare delivery guaranteed thereby allowing the hospital ED to better meet the unpredictable increase in service demand. Future work may consider the use of exploratory data analytics like principal component analysis for data pre-processing and dimensionality reduction.

Conflict of Interest: The authors declare that they have no conflict of interest.

Acknowledgments: The authors would like to thank the Tertiary Education Trust Fund (TETFund), Nigeria for supporting this research through the TETFund Center of Excellence in Computational Intelligence Research and the Management of the University of Uyo for providing a conducive environment for research.

References

Abo-Hamad, W. & Arisha, A. (2013). Simulation-based framework to improve patient experience in an emergency department, *European Journal of Operations Research*, 224(1), 154-166. http://dx.doi.org/10.1016/j.ejor.2012.07.028.

ACEP, (2010). *Meeting the challenge of emergency department overcrowding/boarding.* Washington: American College of Emergency Physicians.

Attai, K., Amannejad, Y., Vahdat Pour, M., Obot, O., & Uzoka, F. M. (2022). A systematic review of applications of machine learning and other soft computing techniques for the diagnosis of tropical diseases. Tropical Medicine and Infectious Disease, 7(12), 398. https://doi.org/10.3390/tropicalmed7120398

Asuquo, D. E., Umoh, U. A., Osang, F. B. & Okokon, E. W. (2020). Performance evaluation of C4.5, random forest and naïve bayes in employee performance and promotion prediction, *African Journal of Management Information System*, 2(4), 41 – 55.

Barnes, M., Hanson, C. &Giraud-Carrier, C. (2018). The case for computational health science, *Journal of Health Informatics Research*, 2, 99-110. https://doi.org/10.1007/s41666-018-0024-y

Bagust, A., Place, M. & Posnett, J.W. (1999). Dynamics of bed use in accommodating emergency admissions: Stochastic simulation model, *BMJ*, 319, 155-158. https://doi: 10.1136/bmj.319.7203.155.

Chonde, S., Parra, C. & Chang, C-J. (2013). Minimizing flow-time and time-to-first-treatment in an emergency department through simulation, In Proceedings of the 2013 Winter Simulation Conference on Simulationand Making

Decisions in a Complex World. https://doi.org/10.1109/WSC.2013.6721612.

Ghanes, K., Diakogiannis, A., Jouini, O., Jemai, Z. & Wargon, M. (2014). Key performance indicators for emergency departments: A survey from an operations management perspective, In Working Paper, Paris: Ecole Centrale.

Gül, M. & Guneri, A. F. (2015). Forecasting patient length of stay in an emergency department by artificial neural networks,*Journal of Aaeronautics and Space Technologies,* 8(2), 43-48. https://doi.org/10.7603/s40690-015-0015-7.

Forero, R., Hillman, K., McCarthy, S., Fatovich, D., Joseph, A. & Richardson, D. (2010). Access block and ED overcrowding: Short report. *EMA - Emergency Medicine Australasia,* 22(2), 119-135. https://doi.org/10.1111/j.1742-6723.2010.01270.x

Geue, C., Lewsey, J., Lorgelly, P., Govan, L., Hart, C. & Briggs, A. (2012). Spoilt for choice: Implications of using alternative methods of costing hospital episode statistics. *Health Economics,* 21(10), 1201-1216. https://doi.org/10.1002/hec.1785.

Graff, L. (1999). Overcrowding in the ED: An international symptom of health care system failure, *American Journal of Emergency Medicine,* 17 (2), 208-209. https://doi.org/10.1016/s0735-6757(99)90064-5.

Hashimoto, F. &Bell, S. (2007). Improving outpatient clinic staffing and scheduling with computer simulation, *Journal of General Internal Medicine,* 182-184. https://doi.org/10.1007/BF02600274.

Hashmi, Z, Abidi, S. & Cheah, Y. N. (2002). An intelligent agent-based knowledge broker for enterprise-wide healthcare knowledge procurement, *In Proceedings of the IEEE Symposium on Computer-Based Medical Systems*, 173-178. https://doi.org/10.1109/CBMS.2002.1011373

Hunter-Zinck, H. S., Peck, J. S., Strout, T. D. & Gaehde, S. A. (2019). Predicting emergency department orders with multilabel machine learning techniques and simulating effects on length of stay, *Journal of the American Medical Informatics Association*, 26(2), 1427-1436. https://doi.org/10.1093/jamia/ocz171.

Magoulas, G.D. & Prentza, A. (2001). Machine learning in medical applications, In Paliouras G, Karkaletsis V, Spyropoulos C, eds. *Machine Learning and Its Applications*, Lecture Notes in Computer Science Series, 2049, 300-307. https://doi.org/10.1007/3-540-44673-7_19.

Manupati, V. K., Teja, G. M., SK Z. Hussain, Sandeep, Y. & Varma, A. A. (2015). Optimisation based on simulation: A patient admission scheduling problem in a radiology department case study, *International Journal on Cybernetics & Informatics*, 4(2), 241-249. https://doi.org/10.5121/ijci.2015.4223 241.

Mariki, M., Mkoba, E., & Mduma, N. (2022). Combining clinical symptoms and patient features for malaria diagnosis: machine learning approach. Applied Artificial Intelligence. 36(1), 2007-2031. https://doi.org/10.1080/08839514.2022.2031826

Moss, J. & Xiao, Y. (2004). Improving operating room coordination: Communication pattern assessment, *The Journal of Nursing Administration*, 93-100. https://doi.org/10.1097/00005110-200402000-00008

Nwosu, A. D., Ossai, E. N., Mba, U. C., Anikwe, I., Ewah, R., Obande, B. O. & Achor, J. U. (2020). Physician

burnout in Nigeria. BMC Health Services Research, 20(1), 1-9.https://doi.org/10.1186/s12913-020-05710-8

Obot, O., Asuquo, D., Attai, K., Johnson, E., Arnold, K., Edoho, M., Ekpenyong, M., Akwaowo, C., Udoh, O., Usen, V. & Uzoka, F-M. (2023). Development of a Mobile App Diagnostic System for Tropical Febrile Diseases, In Piet Kommers, Mario Macedo, Guo Chao Peng, & Ajith Abraham (eds.). Proceedings of the 15th International Conference on e-Health and 17th Multi Conference on Computer Science and Information Systems (MCCSIS 2023), pp. 183-190, 15-18 July, Porto, Portugal.

Oddoye, P., Jones, F., Tamiz, M. & Schmidt, P. (2009). Combining simulation and goal programming for healthcare planning in a medical assessment unit, European Journal of Operational Research 193(1), 250-261. https://doi.org/ DOI:10.1016/j.ejor.2007.10.029

Othman, S. B. & Hammadi, S. (2017). A multi-criteria optimization approach to health care tasks scheduling under resources constraints,*International Journal of Computational Intelligence Systems*, 10 (1), 419-439. https://doi.org/10.2991/ijcis.2017.10.1.29

Poulymenopoulou, M., Malamateniou, F. & Vassilacopoulos, G. (2008). *Emergency healthcare process automation using workflow technology and web services, Medical* Informatics and the Internet in Medicine, 28(3), 195-207. https://doi.org/10.1080/14639230310001617841.

Priya, G., AnandhaKumar. P. & Maheswari, K. (2008). Dynamic scheduler - a pervasive healthcare system insmart hospitals using RFID. *In Proceedings of the International Conference on Computing, Communication and*

Networking, 1-6. https:doi.org/10.1109/ICCCNET.2008.4787683

Rashwan, W., Habib, H., Arisha, A., Courtney, G. & Kennelly, S. (2016). An integrated approach of multi-objective optimization model for evaluating new supporting program in Irish hospitals, In T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, eds., *Proceedings of the IEEE 2016 Winter Simulation Conference*, Washington DC, 11th -14th December, 1904-1915.

R Core Team (2019). R: A language and environment for statistical computing, R foundation for statistical computing, Vienna, Austria. https://www.R-project.org/.

Saxena, R. (2017). How decision tree algorithm works, data aspirant, 30 January 2017. [Online]. Available: https://dataaspirant.com/2017/01/30/howdecision-tree-algorithm-works/ (accessed 26.08.19).

Sch¨olkopf, B. & Smola, A. J. (2002). Learning with kernels, MIT Press, London.

Shafaf, N. & Malek, H. (2019). Applications of machine learning approaches in emergency medicine: A review, *Archives of Academic Emergency Medicine*, 7(1), 34.

Smola, A. J., O´va´ri, Z. L.& Williamson, R. C. (2001). Regularization with dot-product kernels, In Leen T. K., Dietterich, T.G. & Tresp, V. (Eds.), *Advances in Neural Information Processing Systems,* 13, MIT Press, 308-314.

Smola, A. & Sch¨olkopf, B. (2004). A tutorial on support vector regression, *Statistics and Computing*, 14, Kluwer Academic Publishers, Netherland, 199-222.

Taboada, M., Cabrera, E., Iglesias, M. L., Epelde, F. & Luque, E. (2011). An agent-based decision support system for hospitals emergency departments, Procedia Comput Sci., 4, 1870–1879. https://doi.org/10.1016/j.procs.2011.04.203

Thorwarth, M., Rashwan, W. & Arisha, A. (2015). An analytical representation of flexible resource allocation in hospitals, *Flexible Services and Manufacturing Journal*, 1-18. https://doi.org/10.1007/s10696-015-9216-4.

Tomar, D. & Agarwal, S. (2013). A survey on data mining approaches for healthcare, *International Journal of Bio-Science and Bio-Technology*, 5(5), 241–66. https://doi.org/10.14257/ijbsbt.2013.5.5.25

Tsai, C-W., Chiang, M-C., Ksentini, A. &Chen, M. (2016). Metaheuristic algorithms for healthcare: Open issues and challenges, *Computers and Electrical Engineering*, 1-14, http://dx.doi.org/10.1016/j.compeleceng.2016.03.005.

Tsai, C. W. & Rodrigues, J. (2014). Metaheuristic scheduling for cloud: A survey, IEEE Systems Journal, 8(1), 279-291. https://doi.org/10.1109/JSYST.2013.2256731

Umoren, I., Udonyah, K. & Isong, E. (2019). A computational intelligence framework for length of stay prediction in emergency healthcare services department, *14th International Conference on Computer Science & Education (ICCSE)*, 539-549, Toronto, ON, Canada. https://doi.org/10.1109/ICCSE.2019.8845332.

Wang, P.-W. & C.-J. Lin (2014). Support vector machines, data classification: Algorithms and applications, CRC Press, New York, USA, 187-204.

Welch, S., Augustine, J., Camargo, C. A.& Reese, C. (2006). Emergency department performance measures and benchmarking summit, *Academic Emergency Medicine*, 1074-1080.https://doi.org/10.1197/j.aem.2006.05.026

Yousefi, M., Yousefi, M., Ferreira, F. R. M., Kim, J. H. & Fogliatto, F. S. (2018). Chaotic genetic algorithm and adaboost ensemble metamodeling approach for optimum resource planning in emergency departments, *Artificial Intelligence in Medicine*, (84), 23-33. https://doi.org/10.1016/j.artmed.2017.10.002

Yeh, J.Y. & Lin, W.S. (2007). Using simulation technique and genetic algorithm to improve the quality careof hospital emergency department,*Expert Systems with Applications*, 32(4), 1073-1083. https://doi.org/10.1016/j.eswa.2006.02.017

Yoo, I., Alafaireet, P., Marinov, M., Pena-Hernandez, K., Gopidi, R. & Chang, J.F. (2012).Data mining in healthcare and biomedicine: A survey of the literature, *Journal of Medical Systems*, 36(4), 2431-2448. https://doi.org/10.1007/s10916-011-9710-5