OPEN ACCESS
International Journal of Innovations in Science & Technology

Prediction of Elective Patients and Length of Stay in Hospital

Muhammad Shahzad Saleem, Saif Ullah
UET Taxila, Punjab, Pakistan
***Correspondence**: Engrshahzad7@gmail.com

Citation | Saleem. M. S, Ullah. S, "Prediction of Elective Patients and Length of Stay in Hospital", IJIST, Vol. 6 Issue. 2 pp 849-865, June 2024
Received | May 26, 2024 **Revised |** June 19, 2024 **Accepted |** June 23, 2024 **Published |** June 27, 2024.

The efficient management of hospital resources and the optimization of patient care are critical tasks in healthcare systems worldwide. One of the key challenges in hospital management is predicting the duration of a patient's stay and accurately determining their location within the hospital, such as whether they are in the Intensive Care Unit (ICU) or the Operating Theater (OT). In this study, we address this problem statement by employing machine learning algorithms to predict both the stay duration and the location of patients within the hospital. The methods applied in this study include Random Forest, Support Vector Machine (SVM), and K-nearest neighbors (KNN) algorithms. These algorithms utilize patient demographic information such as age, weight, and severity of disease as features to predict the stay duration and location. The dataset used for this study consists of a revised dataset containing relevant patient information. Upon applying the machine learning algorithms, we obtained promising results. The Random Forest algorithm achieved the highest accuracy of 88.6% in predicting patient locations, followed by SVM with an accuracy of 60.8% and KNN with an accuracy of 58.1%. Additionally, Random Forest exhibited superior precision, recall, and F1-scores for both ICU and OT classifications compared to SVM and KNN. The results obtained from this study have several practical implications and potential uses. Firstly, accurate predictions of patient stay duration and location can aid hospital administrators in resource allocation and planning, enabling them to efficiently manage bed occupancy and staffing levels. Additionally, healthcare providers can use these predictions to anticipate patient needs and allocate resources accordingly, thereby enhancing patient care and satisfaction. Moreover, the machine learning algorithms utilized in this study can be integrated into hospital information systems to automate the prediction process, providing real-time insights to healthcare professionals. In conclusion, the application of machine learning algorithms in predicting patient stay duration and location within the hospital offers promising results and valuable insights for hospital management. By leveraging patient demographic information and advanced predictive models, healthcare institutions can improve operational efficiency, enhance patient care delivery, and ultimately optimize resource utilization.

Keywords: Hospital Management; Patient Stay Duration; Patient Location; Machine Learning Algorithms.

Introduction:

To make sure that patients receive the best care, hospitals are paying more attention to how they manage their resources. They want to spend less money while still giving good care to patients. Hospital managers focus on planning by deciding what facilities and staff are needed to run the hospital well and meet patients' needs. There are different methods to predict how many patients will come to the hospital, how many beds will be needed, and how those beds will be used. The most important part of these methods is accurately guessing how long patients will stay in the hospital and understanding what things affect how long they stay. The stay in hospital (StayinHos) for a patient in a hospital refers to the number of days they spend in the hospital during one admission [1]. It's a key factor in understanding how much of the hospital's resources a patient consumes. StayinHos also helps us understand how patients move through different parts of the hospital, which is important for evaluating how well the hospital is running. StayinHos is often used to measure how much resources are being used, the cost of treatment, and how severe a patient's illness is [2][3]. Some studies have tried to group patients based on their medical conditions, assuming that each condition has a recommended stay in the hospital [4] However, StayinHos is a complicated concept influenced by many different factors, some of which may even compete with each other. These factors include a patient's characteristics, reason for admission, any complications that arise during their stay, and discharge plans. Additionally, the complexity of treatment can also impact StayinHos, often leading to longer stays than initially anticipated. Therefore, having a model that accurately predicts a patient's StayinHos during their hospital stay could help healthcare providers take proactive measures to prevent unnecessary extensions of stay. Many patients would prefer to receive care in the comfort of their own homes, especially for services like palliative care, if it's feasible and appropriate [5]. Moreover, there are potential risks for patients who stay in the hospital longer than necessary for active treatment. Remaining in the hospital when they could be discharged increases the chances of experiencing falls, acquiring infections specific to hospitals, and encountering medication errors. These risks are particularly concerning for patients who are fit to leave the hospital. Proactively managing discharge procedures starting early in the admission process and minimizing stay in the hospital can help prevent such complications [6][7]. Reducing the stay in hospital is desirable for healthcare providers for two main reasons. Firstly, it allows them to tailor the level of care to each patient's specific needs [8] Secondly, it contributes to more efficient management and allocation of healthcare resources. By minimizing the time each patient spends in the hospital, resources can be distributed more effectively among a larger number of patients [9][10][11].

Objective:

The objective of the study is to develop and apply machine learning algorithms to predict the duration of a patient's stay and their location within the hospital, such as whether they are in the Intensive Care Unit (ICU) or the Operating Theater (OT). By leveraging patient demographic information, such as age, weight, and severity of disease, the study aims to provide accurate predictions that can assist hospital administrators and healthcare providers in efficiently managing resources, planning bed occupancy, and staffing levels, ultimately enhancing patient care and satisfaction. The study evaluates the performance of different machine learning algorithms, namely Random Forest, Support Vector Machine (SVM), and K-nearest neighbors (KNN), and finds that Random Forest achieves the highest accuracy and superior performance metrics in predicting patient locations. The results suggest practical applications for integrating these predictive models into hospital information systems to provide real-time insights and optimize resource utilization in healthcare institutions.

Novelty:

The novelty of the method described in the study lies in its comprehensive and integrated approach to improving hospital management through machine learning. By employing a comparative analysis of multiple machine learning algorithms, including Random Forest, Support Vector Machine, and K-nearest neighbors, the study evaluates their effectiveness in predicting both the duration of a patient's stay and their specific location within the hospital, such as the ICU or OT. This dual focus on predicting both stay duration and location is particularly innovative, as it provides a more holistic solution to managing hospital resources. The study utilizes patient demographic information, such as age, weight, and severity of disease, as predictive features, highlighting the significant role these factors play in patient outcomes and demonstrating their effective integration into predictive models. Notably, the Random Forest algorithm achieved a high accuracy of 88.6% in predicting patient locations, outperforming SVM and KNN, and exhibited superior precision, recall, and F1 scores for ICU and OT classifications. The practical implications of these findings are substantial, as accurate predictions can aid hospital administrators in resource allocation and planning, optimizing bed occupancy and staffing levels. Furthermore, the potential for real-time integration of these predictive models into hospital information systems represents a significant advancement toward automating and enhancing hospital management processes. Overall, this method promises to improve operational efficiency and patient care delivery by leveraging advanced predictive models and patient demographic information, ultimately optimizing resource utilization in healthcare institutions.

Flow of Study:

Here is the flow of the study and it comprises these steps, namely, Data Collection, Data Preprocessing, Feature Selection, Model Selection, Model Training, Model Evaluation, Comparison and Analysis, Implementation and Integration, and Practical Implications.

Approaches for Stay in Hospital of Patients:

As healthcare costs continue to rise, accurately predicting how long a patient will stay in the hospital due to a serious illness or disease is becoming increasingly crucial for planning and evaluating healthcare interventions [12]. These costs can escalate due to various factors, such as the range of medications and treatments provided, staffing requirements, and the use of specialized equipment, in addition to the length of a patient's stay. Swiftly identifying patients who are at a higher risk of experiencing a prolonged hospital stay or death can help significantly reduce these unavoidable costs. It can also enhance patient care and decrease the likelihood of patients experiencing further healthcare-related complications [13]. A significant portion of research on StayinHos has mainly focused on identifying the factors that strongly impact StayinHos in various settings rather than directly predicting StayinHos outcomes [14][15]. There's been limited exploration using machine learning models that directly address StayinHos. Present machine learning studies typically concentrate on specific patient groups and medical conditions [16][17]. However, there's growing interest in innovative deep-learning methods, especially with the increasing use of Electronic Health Records (EHRs) in healthcare settings [18][19]. This interest has led to the creation of advanced predictive modeling techniques aiming to enhance healthcare quality and provide more personalized care. These models are applied across various clinical prediction tasks, including predicting StayinHos [20][21].

It is noteworthy that a considerable amount of scholarly literature has been dedicated to addressing the challenges associated with predicting both the length of StayinHos and mortality rates within hospital settings. Prediction models for these issues commonly utilize conventional arithmetic methods, such as mean and median calculations, alongside statistical techniques like regression analysis [22]. Furthermore, data mining techniques have emerged as valuable tools in these domains. Hospitals face ongoing pressures to improve patient care quality while simultaneously reducing costs, particularly within intensive care units (ICUs) where care needs are intricate and expenses are heightened. Evaluating care efficiency within ICUs often entails assessing hospital mortality rates and StayinHos durations. Consequently, many of the predictive models developed for StayinHos estimation are equally applicable to forecasting mortality outcomes [23]. This section aims to provide an overview and comprehensive survey of the diverse analytical approaches documented in the literature about StayinHos prediction over recent decades.

Research-Based Techniques for Stay in Hospital of Patients:

A commonly employed metric within the realm of StayinHos prediction is the average StayinHos, often computed straightforwardly by determining the mean value. This average duration of StayinHos is typically derived by dividing the total number of in-patient days by the count of patient admissions sharing the same diagnosis-related group classification. However, it's important to note that this metric may inadequately capture the underlying data distribution, particularly when dealing with highly skewed data [24]. Despite its limitations, the average StayinHos finds widespread use in basic planning and hospital capacity management due to its simplicity. Models formulated based on this metric often adopt deterministic approaches and rely heavily on spreadsheet-based calculations [25]. Nevertheless, given the inherently complex and uncertain nature of hospital environments, simplistic methodologies may prove less effective [26]. Dependence solely on the average StayinHos can be misleading, especially in scenarios where the data distribution deviates from normality. Models exclusively built upon the average StayinHos may fail to accurately represent the patient population, potentially leading to erroneous conclusions [27]. To tackle this challenge, operational strategies have been devised to enhance patient flow modeling and StayinHos prediction. In the ensuing sections, we delve into four operational modeling techniques utilized for StayinHos prediction, elucidating their respective applications.

Computational Modelling

Compartmental systems, as defined by literature [28], consist of a finite number of homogeneous, well-mixed, lumped subsystems referred to as compartments. These models come in various forms, encompassing linear, deterministic, non-linear, or stochastic representations, contingent upon the specific process they aim to emulate. Over recent decades, compartmental models have found application in studying patient movement across different hospital systems. Notably, in a study cited in [29], it was observed that StayinHos durations within geriatric departments couldn't be accurately encapsulated by a singular metric like the mean value. Instead, the StayinHos distribution exhibited characteristics akin to a mixed exponential distribution. Here, an exponential distribution delineates the probability distribution of time intervals between events, while a mixed distribution portrays the probability distribution of a random variable derived from a collection of other random variables [30].

The model described in [31] provides a mechanism for estimating the count of acute and long-stay patients, along with predicting their expected hospital stays. By forecasting both the average StayinHos and the average number of patients in each state, this two-compartment model aids healthcare professionals in effectively managing bed utilization within the geriatric department. It offers valuable insights into patient flow dynamics within the department and forecasts the anticipated StayinHos. Building upon this framework, subsequent research in [31] extended the model's capabilities by introducing a third compartment dedicated to rehabilitative care, thereby enriching the representation of patient flow dynamics within the hospital setting.

Compartmental modeling strategies rely on a daily census, where parameters characterizing fluctuating flow rates are deduced from occupancy profile data collected over a single day [32]. Depending on the presence of one, two, or three components in the best-fit mixed exponential equation, corresponding statistics for one, two, or three compartments are generated. This modeling technique has found successful applications across various healthcare domains, as evidenced by studies such as those cited in [33] and [34]. Notably, in [35], researchers examined a one-night bed occupancy census encompassing data from 6,068 patients across seven distinct provider groups to model a comprehensive health and social care system, encompassing geriatric hospital beds, psychiatry beds, and nursing homes. Compartmental modeling stands out as a well-established and mathematically rigorous methodology consistently employed to simulate patient flow dynamics within healthcare systems. Nonetheless, the predominant focus of compartmental modeling approaches has often been directed toward specific patient cohorts, such as geriatric patients. The inclination towards specializing compartmental modeling primarily for specific patient groups, such as geriatric patients, has impeded the widespread adoption of these models by both researchers delving into StayinHos studies and hospital decision-makers. However, with the recent surge in interest in "big data" and the increasing focus on delving into electronic health records, it is conceivable that the utilization of compartmental modeling might wane, giving way to more contemporary machine learning techniques like artificial neural networks. A notable limitation of compartmental models lies in their reliance on a solitary day's census of beds, rendering them highly sensitive to the specific day on which the census was conducted. Consequently, the StayinHos predictions derived from these models might struggle to generalize well over prolonged durations. Moreover, compartmental models overlook cyclical patterns in admissions and discharges, failing to incorporate seasonal variations in the data.

Simulation Modelling:

Expanding compartmental models can entail adopting a queueing system perspective, which involves utilizing metrics like "time in the system" and "time spent waiting in a queue" to gauge performance. By employing queueing systems or simulations, hospital planners can evaluate and strategize for various scenarios, thereby alleviating bottlenecks within the system. Discrete event simulation (DES) serves as a prominent approach in this regard, wherein the system is modeled as it progresses through time, with state variables undergoing instantaneous changes at discrete events. These events signify occurrences capable of altering the system's state, and DES has garnered extensive usage in modeling healthcare systems. In general, patient simulation systems comprise fundamental components aimed at replicating real-world hospital scenarios:

- **Activities:** These encompass the operations and tasks responsible for transforming entities within the system.
- **Entities:** These represent the elements of the simulated system, such as patients undergoing treatment.
- **Overall State:** This denotes a comprehensive collection of features describing the entire system's current status.

The concept of such systems was initially introduced in a study documented in [36], where a simulation model developed using discrete event simulation (DES) was utilized for numerical evaluations. This simulation model featured three compartments: acute, rehabilitation, and long-stay, which were constructed and validated against results obtained from established compartmental models. A significant finding from this research was the identification of a "long warming up period," during which continuous simulations were conducted until the system reached a steady state. This observation suggested that any modifications made to the system, such as adjusting the number of beds or altering patient StayinHos, required an extended period for the model to stabilize. By manipulating policy parameters like the overall level of emptiness, number of available beds, ward conversion rates, and patient admissions, the simulation model could assist hospital planners in assessing the effectiveness of a geriatric department.

In essence, simulation modeling offers hospital managers not only the ability to test changes within a system but also greater flexibility in comprehending the system under investigation. This comprehension can be further enriched by incorporating external compartments such as independent homes and support homes into the basic model configuration, as suggested in [37]. Furthermore, the basic model can be adapted to address potential scenarios such as a winter bed crisis in English hospitals, as explored in [38]. However, the practical application of these models in real-world settings often necessitates a strong operational focus and the collection of substantial volumes of data. Additionally, simulation models tend to apply only to the specific environments in which they are deployed, lacking generalizability. While they can provide valuable insights, the extensive data requirements and development costs associated with simulation models may pose challenges to their widespread adoption by clinicians.

Markov Modelling:

Markov and semi-Markov models are constructed under the premise that patient subgroups exhibit homogeneity, resulting in events occurring at regular intervals in time. These modeling techniques are instrumental in comprehending patient flow and StayinHos, particularly in larger population cohorts where the assumptions of Markov models hold [39]. Markov models operate on the premise of probabilistic patient behavior within healthcare systems, thus providing a realistic depiction of real-world healthcare dynamics. A continuous time stochastic model delineating patient flow is expounded upon in [40]. This model constitutes a two-stage continuous time Markov model elucidating the progression of geriatric patients through geriatric hospitals. The compartments in the model signify distinct states, and the probabilities governing patient transitions between these states can be quantified. Patients initially admitted to an acute stay state may transition to a long-stay state or depart the hospital entirely through discharge or death states, akin to the compartments in compartmental modeling.

The work delineated in [40] builds upon that of [30], which investigated the StayinHos distribution of patients with a given census date utilizing a mixed exponential distribution. However, the model presented in [40] is deterministic and discrete-time valued, facilitating the estimation of the number of patients admitted to acute and long-stay states and their anticipated StayinHos. In contrast, [40] extends this methodology to continuous time, enabling the computation of variances and covariances for acute and long-stay patients. Two distinct models are formulated: the first assumes a waiting list of patients with a constant overall patient count, while the second model portrays a scenario featuring random admissions. An augmentation of the stochastic Markov model mentioned earlier was devised to accommodate three stages [41]. This extension incorporates a cost element for each stage, allowing the model to support health and social services for the elderly while factoring in associated costs.

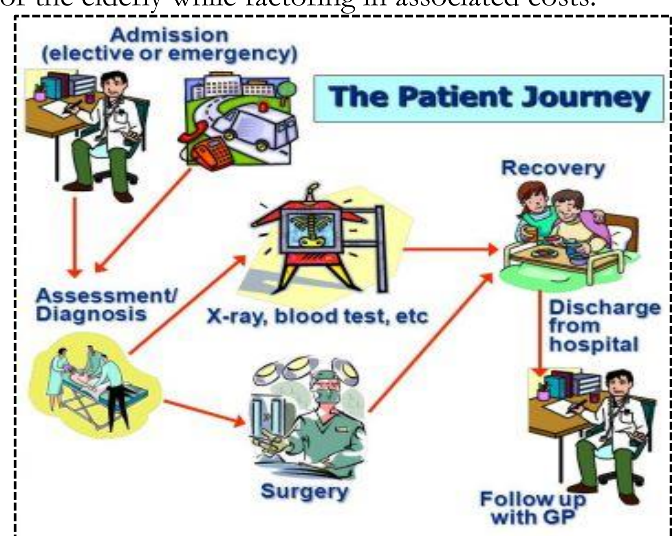

Figure 1: Patient's Journey [42].

Expanding on previous research, [43] employs a comparable methodology to a four-compartment model outlined in [44], where the four stages delineate acute, long-stay, community, or deceased statuses. This model facilitates the estimation of the anticipated patient count in each stage at any given time for multiple patients admitted on the same day. Subsequent enhancements to these models were documented in [45], which introduced six stages aiming to

investigate the intrinsic interactions between hospital medical services and community care. Figure 1 illustrates the trajectory of patients through these stages. The StayinHos in hospitals, particularly in intensive care units (ICUs) and operating theaters, is a critical factor affecting patient outcomes, healthcare resource utilization, and overall hospital efficiency. The severity of diseases often correlates with prolonged hospital stays, posing challenges for healthcare providers in managing patient flow, allocating resources, and optimizing care delivery. Additionally, the dynamic nature of patient conditions and the unpredictable nature of medical emergencies further complicate the process of predicting and managing StayinHos.
The major contributions of our work are given below,

1. We developed an AI-based approach for predicting the stay of patients in hospitals and patients in ICU and OT.
2. The proposed method effectively predicted the stay of patients in hospitals and patients in ICU and OT.
3. Comparative analysis with different AI-based approaches demonstrated the efficient capability of our technique in predicting the stay of patients in hospitals and patients in ICU and OT.

We prepared the remaining paper for the following purposes: Section 2 delves into the proposed working system, Section 3 provides detailed information on the experimental evaluation, and lastly, Section 4 draws conclusions from this work.

Proposed System:

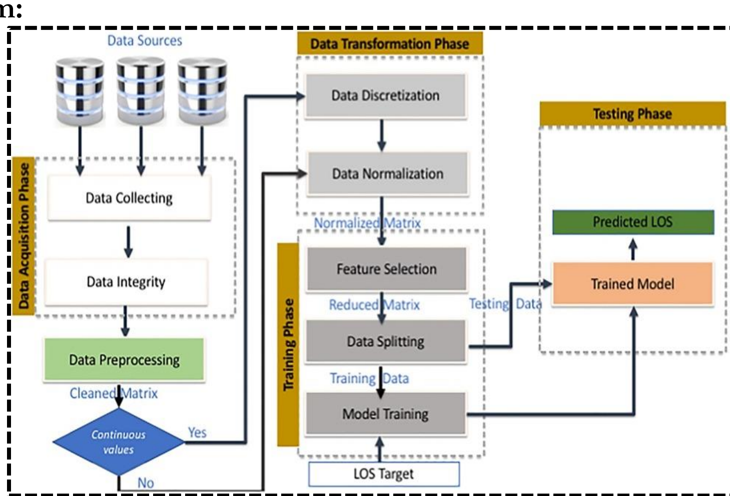

Figure 2: Proposed Working System.

This work aims to predict the stay in hospital of patients and the location such as ICU and OT. The detailed working mechanism is illustrated in Figure 2.

Random Forest Regression:

Decision Trees serve dual purposes, addressing both regression and classification challenges. Structured like branching trees, they visually depict decision paths, hence their name. In regression scenarios, they commence at the tree's root, traversing splits dictated by variable outcomes until reaching a leaf node, where the final result is derived. Figure 3 is an illustrative example of a decision tree.

In this basic decision tree diagram, the process begins with Var_1 and proceeds to split based on predetermined criteria. If the condition is met ('yes'), the tree follows one path; if not ('no'), it takes the alternative route. This cycle continues until the tree reaches a leaf node, where the outcome is determined. In the provided example, the placeholders a, b, c, or d can represent any numeric or categorical values. Ensemble learning involves employing multiple models trained on the same dataset and aggregating their predictions to achieve a more robust predictive or classification outcome. The premise behind ensemble learning is to capitalize on the diversity among models, such as decision trees so that their errors are independent and vary from one model to another.

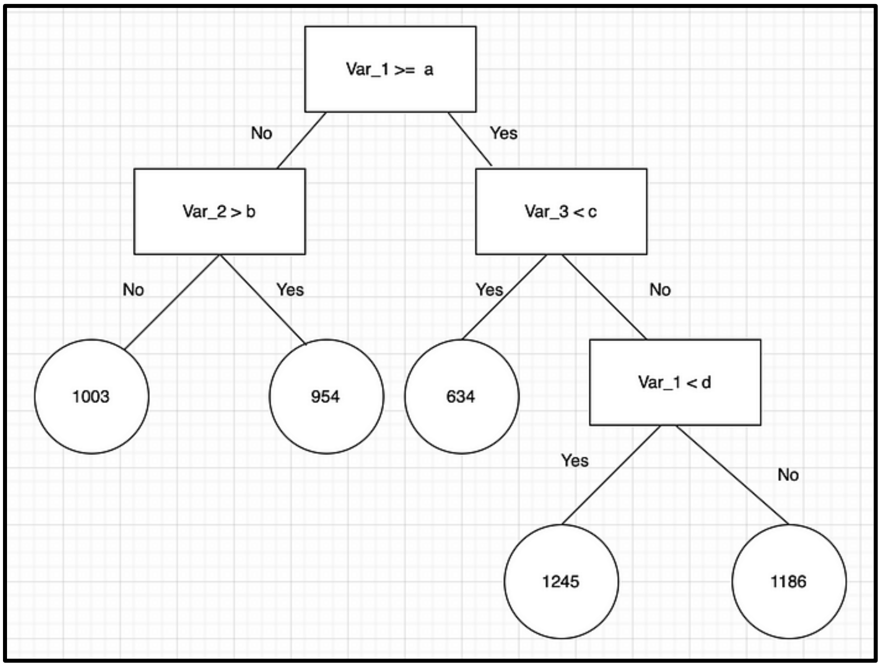

Figure 3: Decision Tree [46]

Bootstrapping refers to the technique of randomly sampling subsets of a dataset across multiple iterations and variable selections. These sampled results are then aggregated or averaged to enhance the robustness of the outcome. Bootstrapping serves as a practical example of an ensemble model in action. The bootstrapping Random Forest algorithm integrates ensemble learning principles with the decision tree framework. It generates multiple decision trees by randomly sampling from the dataset, then combines their predictions through averaging to produce a consolidated result, which typically yields robust predictions or classifications. In the Random Forest Regression Model, the sklearn module was utilized for training our random forest regression model, specifically the Random Forest Regressor function. The Random Forest Regressor documentation showcased many different parameters that are selected for our model. Some of the important parameters are highlighted below:

- **n_estimators**: This parameter specifies the number of decision trees that will be included in the random forest model. Increasing the number of estimators typically leads to better performance, but it also increases computational complexity and training time.
- **criterion**: This parameter allows you to choose the criterion or StayinHoss function used to measure the quality of a split in the decision tree. The two main options are Mean Squared Error (MSE) and Mean Absolute Error (MAE). The default value is MSE, which is often suitable for regression tasks.
- **max_depth**: This parameter controls the maximum depth of each decision tree in the forest. A deeper tree can capture more complex relationships in the data, but it also increases the risk of overfitting. Setting an appropriate value for max_depth helps prevent overfitting and improves generalization to unseen data.
- **max_features**: This parameter specifies the maximum number of features that the model will consider when looking for the best split at each node of the decision tree. By limiting the number of features, you can reduce the computational burden and prevent the model from focusing too much on noisy or irrelevant features.
- **bootstrap**: When set to True (the default), this parameter enables bootstrapping, a resampling technique where random samples are drawn with replacement from the

training dataset. Bootstrapping helps create diverse subsets of data for each tree, contributing to the randomness and robustness of the random forest model.

- **max_samples**: This parameter is relevant only when bootstrap is set to True. It specifies the maximum number of samples (or proportion of samples) to be drawn for training each tree. Controlling the size of the bootstrap samples can influence the diversity of the trees and the overall performance of the random forest.
- **min_samples_split**: This parameter determines the minimum number of samples required to split an internal node in a decision tree. If the number of samples at a node is less than min_samples_split, the node will not be split, and it will become a leaf node. Setting a higher value for min_samples_split can help prevent the model from overfitting by imposing constraints on the minimum number of samples required for a split, thereby promoting simpler and more generalized trees.
- **min_samples_leaf**: This parameter specifies the minimum number of samples required to be at a leaf node. In other words, a split will only be considered if it leaves at least min_samples_leaf training samples in each of the left and right branches. Similar to min_samples_split, setting a higher value for min_samples_leaf helps prevent overfitting by controlling the minimum size of leaf nodes. It encourages the model to create larger, more robust branches.
- **n_jobs**: This parameter controls the number of CPU cores used when fitting the random forest model. By default, n_jobs is set to 1, meaning that the model will only use a single CPU core for training. However, you can set n_jobs to -1 to utilize all available CPU cores, speeding up the training process significantly, especially for large datasets. Using multiple CPU cores allows for parallelization of computations, resulting in faster model fitting and improved efficiency.

Each decision tree in a Random Forest splits the data using the best split among a random subset of the features at each node. For regression, the decision at each node is typically made by minimizing the Mean Squared Error (MSE) or another relevant metric. The MSE for a split is calculated as:

$$\mathrm{MSE} = \frac{1}{N}\sum_{k=0}^{n}\left(\mathrm{yi} - \mathrm{y}^{\mathrm{i}}\right)2 \quad (3.1)$$

Where:

- N is the number of data points in the subset,
- yi is the actual value,
- ỳ ỳỳi is the predicted value for the data point.

For regression, the Random Forest algorithm's prediction (Ỳ) for a given input vector is the average of the predictions from all the individual trees in the forest:

$$\grave{\mathrm{Y}} = \frac{1}{\mathrm{T}}\sum_{t=1}^{\mathrm{T}} \grave{\mathrm{y}}(\mathrm{x}) \quad (3.2)$$

Where:

- T is the total number of trees in the forest,
- ỳ3(x) is the prediction for the input x from the tree.

Dataset Analysis:

The study employed various performance metrics to evaluate the effectiveness of the proposed approach. These included roots mean square error (RMSE), mean square error (MSE), R-squared, accuracy, precision, recall, and F1-score. The subsequent sections delve into a detailed analysis of our experimental results, featuring insights from confusion matrix analysis and performance comparisons.

Dataset Details:

The dataset contains information about patients admitted to a hospital named Surgi Care Hospital, Aziz Bhatti Town Sargodha. It includes the following columns:

- **sr#**: This column represents a serial number or identifier assigned to each record in the dataset. It serves as a unique identifier for referencing individual entries.
- **Patient Number**: This column contains unique identifiers for each patient admitted to the hospital. Each patient is assigned a distinct patient number to differentiate them from others in the dataset.
- **Patient Name**: This column stores the names of the patients admitted to the hospital. It provides information about the identity of each patient.
- **Gender**: This column records the gender of each patient, indicating whether they are male or female. Gender is a categorical variable with two possible values: "Male" or "Female".
- **Age**: This column specifies the age of each patient at the time of admission to the hospital. Age is typically recorded in years and represents a continuous numerical variable.
- **Weight**: This column denotes the weight of each patient upon admission to the hospital. Weight is usually measured in kilograms or pounds and represents a continuous numerical variable.
- **Date of Admission**: This column indicates the date when each patient was admitted to the hospital. It provides temporal information about the timing of admissions.
- **Hospital Name**: This column identifies the name or designation of the hospital where each patient was admitted. It helps in categorizing patients based on the hospital they were treated in.
- **Disease**: This column specifies the disease or medical condition for which each patient was admitted to the hospital. It provides information about the primary reason for hospitalization.
- **Severity of the Disease**: This column quantifies the severity level or extent of the disease or medical condition affecting each patient. It helps in assessing the seriousness of the patient's health condition.
- **Stay in Hospital**: This column indicates the duration of each patient's stay in the hospital following admission. It provides information about the length of hospitalization, typically measured in days or weeks.
- **Location**: This column denotes the specific location within the hospital where each patient is admitted. It categorizes patients based on their placement in different areas of the hospital, such as wards, intensive care units (ICUs), or operating theaters (OTs).

These columns collectively provide comprehensive information about patients admitted to hospitals, including their demographics, medical history, hospitalization details, and current health status. Analyzing this dataset can help in understanding patient demographics, disease patterns, treatment outcomes, and resource utilization within healthcare facilities. The target variable for the first prediction task (stay in hospital prediction) is Stay in Hospital, while the target variable for the second prediction task (location prediction) is Location.

Results and Discussion:

This section has the details of the approaches for the prediction of StayinHos and the location such as ICU or OT. The results for each using different algorithms are discussed in subsequent sections.

Random Forest Regression for StayinHos:

Data Used:

The data utilized in this analysis comprises information about patients' demographics, clinical characteristics, and the duration of their hospital stay. Stored in a CSV file, the dataset contains features such as age, weight, severity of the disease, and the duration of hospitalization. These attributes serve as inputs for predicting the stay duration category. The dataset likely originates from a healthcare setting where patient records are maintained for medical and administrative purposes. By leveraging this data, healthcare providers can gain insights into factors influencing patient length of stay, enabling them to better allocate resources and improve operational efficiency.

Approach:

The approach begins with loading and preprocessing the dataset, involving steps such as converting categorical variables into numerical form using one-hot encoding and splitting the data into training and testing sets. Subsequently, a Random Forest Classifier model is trained using the training data. Once trained, the model is used to predict the stay duration category for the test set. Performance metrics including accuracy, R-squared, Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) are computed to evaluate the model's efficacy. Additionally, visualizations such as confusion matrices, bar plots, and histograms provide further insights into the model's performance and the distribution of predicted values.

Results Obtained:

The analysis yields promising results, indicating that the Random Forest Classifier model accurately predicts the stay duration categories for the patients in the test set. The high accuracy score of 95.375% suggests that the model correctly identifies the stay duration category for the majority of patients. Similarly, the R-squared value of 0.944 indicates a strong correlation between the predicted and actual stay durations. Furthermore, the low MSE and RMSE values of 1.13375 and approximately 1.065, respectively, imply that the model's predictions of StayinHos align with the actual stay durations.

Several factors contribute to the obtained results. Firstly, the Random Forest Classifier is a robust ensemble learning algorithm capable of capturing complex relationships between input features and target variables. Additionally, the dataset likely contains informative features that are highly correlated with the target variable (stay duration), enabling the model to make accurate predictions. Moreover, the preprocessing steps, such as one-hot encoding and splitting the data into training and testing sets, help ensure that the model is trained and evaluated appropriately, leading to reliable results.

Uses of the Method and Results:

The method and results have several practical applications in the healthcare domain. Healthcare providers can leverage the trained model to predict the expected duration of a patient's hospital stay based on their demographic and clinical characteristics. By accurately forecasting stay durations, hospitals can optimize resource allocation, streamline patient flow, and improve overall operational efficiency. Additionally, insights gleaned from the analysis can inform decision-making processes and aid in the development of strategies to enhance patient care, resource management, and hospital operations. Overall, the method and results offer valuable tools for improving healthcare delivery and patient outcomes. The following three figures show the results obtained using the random forest for stay in hospital.

Comparison of All Approaches:

The three methods employed for predicting stay durations in hospitals exhibit varying degrees of performance and predictive accuracy. The Random Forest algorithm yielded the most favorable results, achieving a high accuracy of 95.375%, accompanied by a robust R-squared value of 0.94446, indicating strong predictive capability. Additionally, the Random Forest model demonstrated low mean squared error (MSE) and root mean squared error (RMSE), suggesting precise predictions with minimal deviation from the actual values. In contrast, the Support Vector Regression (SVR) model showed comparatively poorer performance, with a significantly lower R-squared value of 0.04968 and substantially higher MSE and RMSE values, indicating a suboptimal fit to the data and less accurate predictions. Similarly, the K Nearest Neighbors (KNN) algorithm exhibited moderate accuracy at 54.75% but yielded a negative R-squared value and notably high MSE and RMSE values, suggesting limited predictive power and potentially inadequate model fitting. Overall, the Random Forest model emerges as the preferred choice for stay duration prediction due to its superior accuracy and robust performance metrics compared to SVR and KNN. Table 1 shows the performance of all approaches for StayinHos.

Table 1: Performance of All the Approaches for StayinHos.

Method	Results	Key Observations
Random Forest	Accuracy: 0.95375 R-squared: 0.94446 MSE: 1.13375 RMSE: 1.06478	High accuracy and R-squared Low MSE and RMSE
Support Vector	R-squared: 0.04968 MSE: 19.40081 RMSE: 4.40463	Poor predictive performance High MSE and RMSE
K Nearest Neighbors	Accuracy: 0.5475 R-squared: -0.36548 MSE: 27.87625 RMSE: 5.27980	Moderate accuracy Poor predictive performance

Random Forest for Location Prediction:

Data Used:

The data utilized in this method comprises features such as Age, Weight, Disease, Severity of the Disease, and Stay in Hospital, with the target variable being the Location. These features are extracted from the provided dataset stored in a CSV file, containing information relevant to hospital admissions.

Approach:

The approach involves employing a Random Forest Classifier to predict the location of patients within the hospital premises based on the provided features. The dataset is preprocessed by converting categorical variables into dummy/indicator variables to facilitate model training. Subsequently, the dataset is split into training and test sets using the train_test_split function from the sci-kit-learn library. The Random Forest Classifier is then initialized, trained on the training data, and used to predict the location of patients in the test set. Model evaluation is performed using accuracy score, confusion matrix, and classification report metrics.

Results Obtained:

The Random Forest Classifier achieved an accuracy of 88.625% in predicting the location of patients within the hospital. The classification report provides additional insights into the model's performance, indicating high precision, recall, and F1-score for both ICU and OT locations. The confusion matrix visually represents the model's performance, showcasing the number of correct and incorrect predictions for each class. The reported accuracy of 88.625% suggests that the Random Forest Classifier model performed reasonably well in predicting the location of patients within the hospital. This means that the model accurately classified approximately 89 out of every 100 instances. Looking at the classification report, we observe that the model performed better in identifying the 'OT' (Operating Theater) locations, with a precision of 92%, meaning that when it predicted a patient to be in the OT, it was correct 92% of the time. The recall for 'OT' is also high at 89%, indicating that the model correctly identified 89% of all OT instances. Similarly, for the 'ICU' (Intensive Care Unit), the precision and recall values are 83% and 88%, respectively. The F1-score, which is the harmonic mean of precision and recall, is 0.86 for 'ICU' and 0.90 for 'OT'. The weighted average F1-score of 0.89 indicates a good overall balance between precision and recall across both classes. These results suggest that the model effectively captures the underlying patterns in the data and can reliably predict the location of patients based on the provided features. The macro average precision, recall, and F1-score are all around 0.88, indicating a balanced performance across both classes. Overall, the model's ability to accurately classify patient locations can be valuable for hospital management in optimizing resource allocation and improving patient care efficiency.

The high accuracy and performance metrics obtained can be attributed to the effectiveness of the Random Forest Classifier in handling complex datasets with multiple features. Random Forests are known for their ability to handle both categorical and numerical data, handle missing values, and mitigate overfitting. Additionally, the preprocessing steps, including converting categorical variables and splitting the dataset into training and test sets, contribute to the model's robustness and generalization capabilities.

Uses of the Method and Results:

The Random Forest Classifier for location prediction in hospital settings has several practical applications. It can assist hospital administrators and staff in efficiently managing patient flow, resource allocation, and bed occupancy. By accurately predicting patient locations, hospitals can optimize staff deployment, streamline logistics, and ensure timely care delivery. Moreover, the insights gained from the classification report can inform decision-making processes, helping healthcare providers improve operational efficiency and patient outcomes. The method's high accuracy and comprehensive evaluation metrics make it a valuable tool for enhancing hospital management and patient care practices.

Comparison of Approaches for Location Prediction:

Table 2 presents a detailed comparison of the performance metrics for three different machine learning algorithms - Random Forest, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN) - in predicting the location of patients within a hospital. Starting with accuracy, Random Forest demonstrates the highest accuracy of 88.6%, indicating its ability to correctly classify patient locations. In contrast, SVM and KNN exhibit lower accuracies at 60.8% and 58.1%, respectively. Moving to precision, Random Forest achieves a precision of 0.83 for ICU and 0.92 for OT, suggesting that when it predicts a patient to be in the ICU or OT, it is correct approximately 83% and 92% of the time, respectively. On the other hand, KNN shows lower precision for ICU (0.46) compared to Random Forest, indicating a higher rate of false positives in classifying patients in the ICU. Regarding recall, Random Forest demonstrates balanced performance for both ICU and OT with recall scores of 0.88 and 0.89, respectively. However, SVM exhibits a recall of 0.00 for ICU, indicating its failure to correctly identify ICU instances, while achieving a perfect recall of 1.00 for OT. Moving to the F1-score, Random Forest shows the highest scores for both ICU (0.86) and OT (0.90), indicating a good balance between precision and recall. In contrast, SVM achieves a low F1-score of 0.00 for ICU due to the absence of true positives, while KNN exhibits moderate F1-scores for both ICU (0.45) and OT (0.66). Analyzing the macro and weighted average F1-scores, Random Forest outperforms both SVM and KNN, demonstrating superior overall performance in capturing the trade-off between precision and recall across both classes. In summary, Random Forest emerges as the most effective algorithm among the three in predicting patient locations within the hospital, offering a high level of accuracy, precision, recall, and F1 scores for both ICU and OT.

Table 2: Performance for Location Prediction of All Approaches.

Metric	Random Forest	SVM	KNN
Accuracy	0.886	0.608	0.581
Precision (ICU)	0.83	-	0.46
Recall (ICU)	0.88	0.00	0.44
F1-score (ICU)	0.86	0.00	0.45
Precision (OT)	0.92	0.61	0.65
Recall (OT)	0.89	1.00	0.67
F1-score (OT)	0.90	0.76	0.66
Macro Avg F1-score	0.88	0.38	0.56
Weighted Avg F1-score	0.89	0.46	0.58

Conclusion:

In this study, we investigated the application of machine learning algorithms for predicting patient stay duration and location within a hospital setting, specifically focusing on the Intensive Care Unit (ICU) and the Operating Theater (OT). Hospital management faces significant challenges in optimizing resource allocation, staffing levels, and patient care efficiency. Accurate predictions of patient stay duration and location can greatly assist in addressing these challenges and improving overall hospital operations. Throughout this study, we employed three machine learning algorithms: Random Forest, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN). These algorithms were trained using a revised dataset containing relevant patient demographic information such as age, weight, disease severity, and stay duration. The primary goal was to evaluate the performance of these algorithms in predicting patient location accurately. Our analysis revealed several key findings. Firstly, the Random Forest algorithm achieved the highest accuracy of 88.6% in predicting patient locations, followed by SVM with an accuracy of 60.8% and KNN with an accuracy of 58.1%. These results suggest that Random Forest outperforms SVM and KNN in accurately classifying patients into the ICU and OT categories based on their demographic attributes. Furthermore, we assessed the precision, recall, and F1 scores for both ICU and OT classifications across the three algorithms. Random Forest exhibited superior performance in terms of precision, recall, and F1-scores for both ICU and OT categories compared to SVM and KNN. This indicates that Random Forest can effectively balance precision and recall in classifying patients into different locations within the hospital. The results obtained from this study have significant implications for hospital management and patient care. Accurate predictions of patient location can enable hospital administrators to optimize resource allocation, streamline bed management, and enhance patient flow. By anticipating patient needs and allocating resources accordingly, hospitals can improve operational efficiency and reduce wait times, leading to better patient outcomes and satisfaction. Moreover, the machine learning algorithms utilized in this study can be integrated into hospital information systems to automate the prediction process. Real-time insights provided by these algorithms can empower healthcare professionals to make informed decisions regarding patient care and resource allocation. Additionally, the predictive models developed in this study can be continuously updated with new data to improve their accuracy and reliability over time. In conclusion, the application of machine learning algorithms for predicting patient stay duration and location offers significant potential benefits for hospital management and patient care. By leveraging patient demographic information and advanced predictive models, healthcare institutions can optimize resource utilization, improve operational efficiency, and enhance the overall quality of care delivery. As the healthcare landscape continues to evolve, machine learning-based approaches will play an increasingly important role in shaping the future of hospital management and patient care.

References:

[1] D. A. Huntley, D. W. Cho, J. Christman, and J. G. Csernansky, "Predicting length of stay in an acute psychiatric hospital," Psychiatr. Serv., vol. 49, no. 8, pp. 1049–1053, Aug. 1998, doi: 10.1176/PS.49.8.1049/ASSET/IMAGES/LARGE/B511T4.JPEG.

[2] A. Lim and P. Tongkumchum, "Methods for Analyzing Hospital Length of Stay with Application to Inpatients Dying in Southern Thailand," Glob. J. Health Sci., vol. 1, no. 1, p. p27, Mar. 2009, doi: 10.5539/GJHS.V1N1P27.

[3] K. C. Chang, M. C. Tseng, H. H. Weng, Y. H. Lin, C. W. Liou, and T. Y. Tan, "Prediction of Length of Stay of First-Ever Ischemic Stroke," Stroke, vol. 33, no. 11, pp. 2670–2674, Nov. 2002, doi: 10.1161/01.STR.0000034396.68980.39.

[4] S. Shea, R. V. Sideli, W. Dumouchel, G. Pulver, R. R. Arons, and P. D. Clayton, "Computer-generated Informational Messages Directed to Physicians: Effect on Length of Hospital Stay," J. Am. Med. Informatics Assoc., vol. 2, no. 1, pp. 58–64, Jan. 1995, doi: 10.1136/JAMIA.1995.95202549.

[5] V. J. Wheatley and J. I. Baker, "'Please, I want to go home': ethical issues raised when considering the choice of place of care in palliative care," Postgrad. Med. J., vol. 83, no. 984, pp. 643–648, Oct. 2007, doi: 10.1136/PGMJ.2007.058487.

[6] T. L. Rodziewicz and J. E. Hipskind, "Medical error prevention," StatPearls [Internet]. StatPearls Publ., 2019.

[7] F. M. Simmons, "CEU: Hospital overcrowding: An opportunity for case managers," Case Manager, vol. 16, no. 4, pp. 52–54, Jul. 2005, doi: 10.1016/J.CASEMGR.2005.06.004.

[8] A. Scott, A. Shiell, and M. G. Farnworth, "The value of early discharge: Dispelling some myths," Health Policy (New. York)., vol. 26, no. 2, pp. 81–91, Dec. 1993, doi: 10.1016/0168-8510(93)90111-2.

[9] A. Clarke, "Why are we trying to reduce the length of stay? Evaluation of the costs and benefits of reducing time in hospital must start from the objectives that govern change.," BMJ Qual. Saf., vol. 5, no. 3, pp. 172–179, Sep. 1996, doi: 10.1136/QSHC.5.3.172.

[10] D. Moher et al., "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement," PLoS Med., vol. 6, no. 7, Jul. 2009, doi: 10.1371/JOURNAL.PMED.1000097.

[11] J. A. Hayden, D. A. van der Windt, J. L. Cartwright, P. Côté, and C. Bombardier, "Assessing bias in studies of prognostic factors," Ann. Intern. Med., vol. 158, no. 4, pp. 280–286, Feb. 2013, doi: 10.7326/0003-4819-158-4-201302190-00009/SUPPL_FILE/AIME201302190-00009_SUPPL1.XLSX.

[12] M. Bauer, L. Fitzgerald, E. Haesler, and M. Manfrin, "Hospital discharge planning for frail older people and their family. Are we delivering best practice? A review of the evidence," J. Clin. Nurs., vol. 18, no. 18, pp. 2539–2546, Sep. 2009, doi: 10.1111/J.1365-2702.2008.02685.X.

[13] M. S. Ullah, H. Qayoom, and F. Hassan, "Viral Pneumonia Detection Using Modified GoogleNet Through Lung X-rays," 2021 4th Int. Symp. Adv. Electr. Commun. Technol. ISAECT 2021, 2021, doi: 10.1109/ISAECT53699.2021.9668553.

[14] P. Yoon, I. Steiner, and G. Reinhardt, "Analysis of factors influencing length of stay in the emergency department," Can. J. Emerg. Med., vol. 5, no. 3, pp. 155–161, 2003, doi: 10.1017/S1481803500006539.

[15] J. E. Mezzich and G. A. Coffman, "Factors Influencing Length of Hospital Stay," https://doi.org/10.1176/ps.36.12.1262, vol. 36, no. 12, Apr. 2006, doi: 10.1176/PS.36.12.1262.

[16] C. S. Yang, C. P. Wei, C. C. Yuan, and J. Y. Schoung, "Predicting the length of hospital stay of burn patients: Comparisons of prediction accuracy among different clinical stages," Decis. Support Syst., vol. 50, no. 1, pp. 325–335, Dec. 2010, doi: 10.1016/J.DSS.2010.09.001.

[17] R. Stoean, C. Stoean, A. Sandita, D. Ciobanu, and C. Mesina, "Ensemble of Classifiers for Length of Stay Prediction in Colorectal Cancer," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 9094, pp. 444–457, 2015, doi: 10.1007/978-3-319-19258-1_37.

[18] B. Shickel, P. J. Tighe, A. Bihorac, and P. Rashidi, "Deep EHR: A Survey of Recent Advances in Deep Learning Techniques for Electronic Health Record (EHR) Analysis," IEEE J. Biomed. Heal. Informatics, vol. 22, no. 5, pp. 1589–1604, Sep. 2018, doi: 10.1109/JBHI.2017.2767063.

[19] P. Yadav, M. Steinbach, V. Kumar, and G. Simon, "Mining electronic health records (EHRs): A survey," ACM Comput. Surv., vol. 50, no. 6, Jan. 2018, doi: 10.1145/3127881/SUPPL_FILE/YADAV.ZIP.

[20] A. Rajkomar et al., "Scalable and accurate deep learning with electronic health records," npj Digit. Med. 2018 11, vol. 1, no. 1, pp. 1–10, May 2018, doi: 10.1038/s41746-018-0029-1.

[21] H. Harutyunyan, H. Khachatrian, D. C. Kale, G. Ver Steeg, and A. Galstyan, "Multitask

learning and benchmarking with clinical time series data," Sci. Data 2019 61, vol. 6, no. 1, pp. 1–18, Mar. 2017, doi: 10.1038/s41597-019-0103-9.
[22] G. Qadir, S. Zareen, F. Hassan, and A. U. Rahman, "Voice Spoofing Countermeasure Based on Spectral Features to Detect Synthetic Attacks Through LSTM," Int. J. Innov. Sci. Technol., vol. 3, no. 4, pp. 153–165, 2021, Accessed: Jun. 24, 2024. [Online]. Available: https://journal.50sea.com/index.php/IJIST/article/view/124
[23] A. Awad, M. Bader–El–Den, and J. McNicholas, "Patient length of stay and mortality prediction: A survey," http://dx.doi.org/10.1177/0951484817696212, vol. 30, no. 2, pp. 105–120, Mar. 2017, doi: 10.1177/0951484817696212.
[24] "Melanoma Detection Using a Deep Learning Approach | International Journal of Innovations in Science & Technology." Accessed: Jun. 24, 2024. [Online]. Available: https://journal.50sea.com/index.php/IJIST/article/view/191
[25] "Modelling for the Planning and Management of Bed Capacities in Hospitals on JSTOR." Accessed: Jun. 24, 2024. [Online]. Available: https://www.jstor.org/stable/822874
[26] P. R. Harper, "A framework for operational modelling of hospital resources," Health Care Manag. Sci., vol. 5, no. 3, pp. 165–173, 2002, doi: 10.1023/A:1019767900627/METRICS.
[27] A. X. Costa, S. A. Ridley, A. K. Shahani, P. R. Harper, V. De Senna, and M. S. Nielsen, "Mathematical modelling and simulation for planning critical care capacity*," Anaesthesia, vol. 58, no. 4, pp. 320–327, Apr. 2003, doi: 10.1046/J.1365-2044.2003.03042.X.
[28] K. Godfrey, "Compartmental models and their application", [Online]. Available: https://tauruspet.med.yale.edu/staff/edm42/courses/ENAS_915_2013/book-chapters/Godfrey_ch2.pdf
[29] P. H. Millard, "Geriatric medicine: a new method of measuring bed usage and a theory for planning, MD Thesis St," Georg. Hosp. Med. Sch. Univ. London, 1988.
[30] G. W. Harrison and P. H. Millard, "Balancing acute and long-term care: The mathematics of throughput in departments of geriatric medicine," Methods Inf. Med., vol. 30, no. 3, pp. 221–228, 1991, doi: 10.1055/S-0038-1634832/ID/BR1634832-13/BIB.
[31] H. G.W., "Compartmental models of hospital patient occupancy patterns. Modelling hospital resource use: a different approach to the planning and control of health care systems," R. Soc. Med., pp. 53–61, 1994.
[32] M. D. Mackay, M., Lee, "Using Compartmental Models to Predict Hospital Bed Occupancy," Dep. Psychol. Univ. Adelaide, pp. 1–18, 2007.
[33] S. I. Millard, P. H., McClean, "Go with the flow: a systems approach to healthcare planning," (Royal Soc. Med. Press Limited, London), 1996.
[34] J. A. García-Navarro and W. A. Thompson, "Analysis of bed usage and occupancy following the introduction of geriatric rehabilitative care in a hospital in Huesca, Spain," Health Care Manag. Sci., vol. 4, no. 1, pp. 63–66, 2001, doi: 10.1023/A:1009657817365/METRICS.
[35] A. Banjar, H. Dawood, A. Javed, and F. Hassan, "Fall event detection using the mean absolute deviated local ternary patterns and BiLSTM," Appl. Acoust., vol. 192, p. 108725, Apr. 2022, doi: 10.1016/J.APACOUST.2022.108725.
[36] F. Hassan, S. Albahli, A. Javed, and A. Irtaza, "A Robust Framework for Epidemic Analysis, Prediction and Detection of COVID-19," Front. Public Heal., vol. 10, p. 805086, May 2022, doi: 10.3389/FPUBH.2022.805086/BIBTEX.
[37] E. Ei-Darzi, C. Vasilakis, T. J. Chaussalet, and P. H. Millard, "A Simulation Model to Evaluate the Interaction between Acute, Rehabilitation, Long Stay Care and the

Community," pp. 475–485, 2000, doi: 10.1007/978-1-4757-4919-9_31.
[38] C. Vasilakis and E. El-Darzi, "A simulation study of the winter bed crisis," Health Care Manag. Sci., vol. 4, no. 1, pp. 31–36, 2001, doi: 10.1023/A:1009649615548/METRICS.
[39] R. Davies and H. Davies, "Modelling patient flows and resource provision in health systems," Omega, vol. 22, no. 2, pp. 123–131, Mar. 1994, doi: 10.1016/0305-0483(94)90073-6.
[40] "Comparative Analysis of Machine Learning Algorithms for Classification of Environmental Sounds and Fall Detection | International Journal of Innovations in Science & Technology." Accessed: Jun. 24, 2024. [Online]. Available: https://journal.50sea.com/index.php/IJIST/article/view/188
[41] S. I. McClean, B. McAlea, and P. H. Millard, "Using a markov reward model to estimate spend- down costs for a geriatric department," J. Oper. Res. Soc., vol. 49, no. 10, pp. 1021–1025, Sep. 1998, doi: 10.1057/PALGRAVE.JORS.2600619/METRICS.
[42] G. Kimberly Sabsay (LSI, "Learning from our past: Using medical history to guide patient care", [Online]. Available: https://insights.princeton.edu/2021/01/analytics-patient-care/
[43] M. P. Taylor G., McClean S., "Continuous-time Markov models for geriatric patient behaviour," Appl. Stoch. Model. Bus. Ind., vol. 13, no. (3-4), pp. 315–323, 1997.
[44] Taylor, G., McClean, S.I., Millard, P.H., "A model to analyse the flow of elderly patients around departments of geriatric medicine and the community. Go with the Flow: A Systems Approach to Healthcare Planning," 1996.
[45] G. J. Taylor, S. I. McClean, and P. H. Millard, "Stochastic Models of Geriatric Patient Bed Occupancy Behaviour," J. R. Stat. Soc. Ser. A Stat. Soc., vol. 163, no. 1, pp. 39–48, Jan. 2000, doi: 10.1111/1467-985X.00155.
[46] "Decision Trees for Classification: A Machine Learning Algorithm", [Online]. Available: https://www.xoriant.com/blog/decision-trees-for-classification-a-machine-learning-algorithm#:~:text=Introduction Decision Trees are a,namely decision nodes and leaves.