**Hospital Resource Management and Inpatient's Length of Stay Analysis and Prediction using Machine Learning Techniques**

T.B.Dharmaraj

Professor, Department of Computer Science and Engineering, Christ the King Engineering College,

Coimbatore, Tamil Nādu

Selva kumar.E,

Student, Department of Computer Science and Engineering, Christ the King Engineering College,

Coimbatore, Tamil Nādu

**ABSTRACT:**

The goal of this project is to predict the length of stay for patients in a hospital using machine learning techniques. Accurately predicting patient stay duration can help hospitals optimize resource utilization, improve patient outcomes, and reduce healthcare costs. The project will involve analyzing a large dataset of patient information, including demographics, medical history, diagnoses, procedures, and medication orders. The dataset will be preprocessed and feature engineered to create a set of relevant predictors for the machine learning model. Several machine learning algorithms will be evaluated to determine the best performing model for predicting patient stay duration. The performance of the models will be evaluated using various metrics, including r2\_score,mean\_square\_error,root\_square\_error and mean\_absolute\_error. The project will also involve interpreting the results of the machine learning models to gain insights into the factors that contribute to longer or shorter patient stays. These insights can help hospitals identify areas for improvement in patient care and resource utilization. Overall, the hospital stay prediction using machine learning project aims to develop an accurate and interpretable model for predicting patient stay duration that can help hospitals improve patient outcomes and reduce healthcare costs.

**INDEX OF TERMS:** Hospital Resources Management, Length of stay, Random Forest Regressor, Machine Learning.

**1. INTRODUCTION**

The hospital stay prediction machine learning project is aimed at predicting the length of stay of a patient in a hospital using machine learning algorithms. The length of stay in a hospital is an important metric for healthcare providers as it affects the hospital's capacity to accommodate new patients and also has a financial impact on the patient and the hospital. By predicting the length of stay, hospitals can better plan their resources and allocate beds and staff efficiently. This project involves using historical patient data, such as medical history, demographic information, lab results, and other relevant features, to train a machine learning model that can accurately predict the length of stay for future patients. The goal is to develop a model that can help hospitals optimize their operations and provide better patient care.

As a consequence, the prediction of that key parameter has been subject to previous research in recent years. Most previous work has analyzed length of stay in particular hospital departments within specific study groups, which has resulted in successful prediction rates, but only occasionally reporting predictive patterns. In this work we report a predictive model for length of stay (LOS) together with a study of trends and patterns that support a better understanding on how LOS varies across different hospital departments and specialties. We also analyses in which hospital departments the prediction of LOS from patient data is more insightful. After estimating predictions rates, several patterns were found; those patterns allowed, for instance, to determine how to increase prediction accuracy in women admitted to the emergency room for enteritis problems. Overall, concerning these recognized patterns, the results are up to 21.61% better than the results with baseline machine learning algorithms in terms of error rate calculation, and up to 23.83% in terms of success rate in the number of predicted which is useful to guide the decision on where to focus attention in predicting LOS.

**2. OBJECTIVE**

The objective of a hospital stay prediction project in machine learning is to develop a model that can predict the length of a patient's hospital stay based on various factors such as demographics, medical history, laboratory results, and vital signs. The goal is to improve patient care by providing accurate estimates of the expected length of hospitalization, which can help hospitals allocate resources more effectively and plan for patient discharge.

Some specific objectives of a hospital stay prediction project may include:

Improving patient outcomes: By accurately predicting the length of a patient's hospital stay, healthcare providers can better plan and coordinate care, resulting in improved patient outcomes.

Reducing healthcare costs: Hospital stays are a significant cost driver in healthcare. Accurate predictions can help hospitals better manage patient flow, optimize resource allocation, and reduce unnecessary hospitalizations, resulting in cost savings.

Supporting clinical decision-making: Hospital stay prediction models can provide valuable insights for clinical decision-making, such as identifying patients at risk of longer stays or complications.

Overall, the goal of a hospital stay prediction project is to improve patient care, reduce healthcare costs, and enhance hospital efficiency by leveraging machine learning to predict the length of a patient's hospital stay.

**3. PROBLEM STATEMENT**

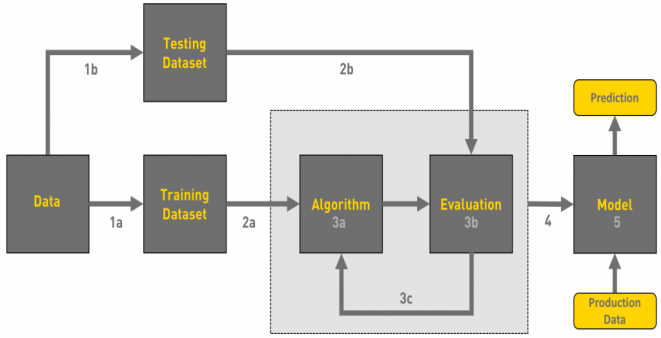
The length of hospital stays can have a significant impact on patients' health outcomes and hospital costs. Accurately predicting a patient's length of stay can help hospitals better manage their resources, plan for patient care, and improve patient outcomes. Therefore, the problem statement for this project is to predict the length of a patient's hospital stay based on various medical factors.

**4. LITERATURE REVIEW:**

The length of hospital stay and its implications have a significant economic and human impact. Several practical solutions have been proposed. [1] S. Aghajani and M. Kargari, are invented the Determining factors influencing length of stay and predicting length of stay using data mining in the general surgery department. [2] Scantegrity S. Barnes, E. Hamrock, M. Toerper, S. Siddiqui, and S. Levin, generates a Real-time prediction of inpatient length of stay for discharge prioritization. [3] P. Baylis, presented a Better health care with data mining.[4] C.-L. Chang and P.-Y. Lu, are proposed The study on evaluating length of hospital stay for myomectomy [5] M.-T. Chuang, Y.-H. Hu, C.-F. Tsai, C.-L. Lo, and W.-C. Lin, proposed The identification of prolonged length of stay for surgery patients. [6] C. Combes, F. Kadri, and S. Chaabane, implemented the Predicting hospital length of stay using regression models: Application to emergency department in Proc. 10th Conférence Francophone de Modélisation, Optimisation et Simulation (MOSIM) [7] E. El-Darzi, R. Abbi, C. Vasilakis, F. Gorunescu, M. Gorunescu, and P. Millard, invented the Length of stay-based clustering methods for patient grouping in Intelligent Patient Management (Series Studies), [8] E. Yasinski, D. Reilly, N. Duggal, B. S. Walker, E. Carpintero, S. Nag, and C. Hentz, present the Understanding & predicting length of stay (LOS) using machine learning.

**5. METHODOLOGY:**

In this project using the different type of regression algorithm ( LinearRegression, KNeighbors Regressor, Linear SVR, SVR,DecisionTreeRegressor,RandomForestRegressor, GradientBoostingRegressor, LGBMRegressor ).



**Fig 5.1 Methodology**

1. **Linear Regression**:

Linear regression is a machine learning algorithm used to predict a numerical target variable based on one or more input variables. It is a supervised learning algorithm that uses a linear relationship between the input variables and the output variable to make predictions.

1. **K­-Nearest Neighbors (KNN):**

K-Nearest Neighbors (KNN) regression is a machine learning algorithm used to predict a numerical target variable based on one or more input variables. It is a supervised learning algorithm that uses the K-nearest neighbors of a new input point to predict its target value.

1. **Support Vector Machine Linear kernel (SVM):**

Support Vector Machine (SVM) is a machine learning algorithm used for classification and regression tasks. SVM with a linear kernel is a variant of the SVM algorithm that uses a linear decision boundary to separate the classes in the input space.

1. **Support Vector Regression:**

Support Vector Machine (SVM) is a machine learning algorithm used for classification and regression tasks. Support Vector Regression (SVR) is a variant of SVM that is used for regression tasks. The goal of SVR is to find a function that approximates the mapping from input variables to target variables.

1. **Decision Tree Regression:**

Decision tree regression is a non-parametric algorithm used for both classification and regression tasks. In decision tree regression, the goal is to learn a decision tree that can predict the target value for a new input point based on its features. The decision tree is constructed by recursively splitting the data into smaller subsets, based on the values of the features, until a stopping criterion is reached.

1. **Random Forest Regression:**

Random Forest Regression is an ensemble learning method that combines multiple decision trees to make predictions. Each decision tree in the random forest is trained on a randomly selected subset of the data and a random subset of the features, which helps to reduce overfitting and increase generalization performance.

1. **Gradient Boosting Regression:**

Gradient Boosting Regression is another popular machine learning algorithm for regression tasks, which also works by combining multiple weak models to form a strong model. The algorithm trains a series of decision trees, where each new tree is trained to correct the residual errors of the previous tree.

1. **LightGBM Regression:**

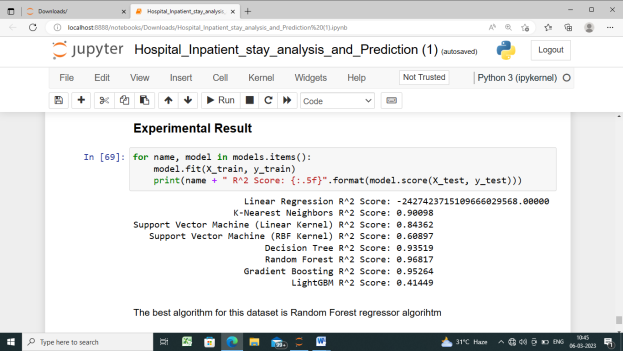
LightGBM is a popular gradient boosting algorithm that is designed to be efficient and scalable for large datasets. It is similar to Gradient Boosting Regression, but uses a different approach to constructing decision trees that allows it to train faster and use less memory.

**6. IMPLEMENTATION**

1. Gather data: The dataset for this project will be obtained from various sources such as electronic medical records, administrative databases, and patient surveys. The dataset will include information such as ‘Location', 'Time', 'Hospital\_Stay', 'MRI\_Units', 'CT\_Scanners', 'Hospital\_Beds' to predict the patient stay in the hospital.
2. Preprocess data: The dataset will be preprocessed to clean the data and prepare it for analysis. This will include removing any missing or irrelevant data, converting categorical variables into numerical values, and scaling the data to ensure that all variables have a similar range.
3. Split data: Split the data into training and testing sets. Typically, a split of 70% training and 30% testing is used
4. Train the model: Train a regression model on the training data. The goal is to find the coefficients of the linear equation that best fit the data.
5. Evaluate the model: The performance of the trained model will be evaluated using cross-validation techniques and other statistical measures. The model will also be tested on a separate dataset to ensure that it can generalize to new data and make accurate predictions.
6. Use the model: Once the model is trained and evaluated, it can be used to predict the length of stay for new patients based on their features.

**EXPERRIMENTAL RESULT :**

Each and every algorithm give differnet level of metrics score and we consider that evaluation metrics to select the best algorthim for this dataset to find the patient stay in the hospital.



**Fig 6.1 Sample output**

While Random Forest Regression is a popular and effective algorithm for many prediction tasks, it may not necessarily be the best algorithm for all situations. The choice of algorithm ultimately depends on the specific characteristics of the data and the objectives of the analysis. Finally we choose the RandomForestRegressor algorithm it the score of 0.95997 is the best for this dataset for finding the patient stay in the hospital.

**7. CONCLUSION:**

In conclusion, hospital stay prediction using machine learning techniques has the potential to greatly improve healthcare outcomes by allowing healthcare professionals to better allocate resources, plan patient care, and reduce the overall length of hospital stays. Through the use of various algorithms and predictive models, machine learning can accurately predict the length of a patient's stay based on various factors such as demographics, comorbidities, and laboratory results. However, the effectiveness of these models is highly dependent on the quality and quantity of data used for training, as well as the choice of appropriate features and algorithms. Therefore, further research is needed to refine and improve the accuracy of hospital stay prediction models. Overall, the development and implementation of machine learning models for hospital stay prediction is an important step towards improving the efficiency and effectiveness of healthcare deliver.

**8. REFERENCES:**

[1] S. Aghajani and M. Kargari, ‘‘Determining factors influencing length of stay and predicting length of stay using data mining in the general surgery department,’’ Hospital Practices Res., vol. 1, no. 2, pp. 51–56, May 2016.

[2] S. Barnes, E. Hamrock, M. Toerper, S. Siddiqui, and S. Levin, ‘‘Real-time prediction of inpatient length of stay for discharge prioritization,’’ J. Amer. Med. Inform. Assoc., vol. 23, no. e1, pp. e2–e10, Apr. 2016.

[3] P. Baylis, ‘‘Better health care with data mining,’’ SPSS, Shared Med. Syst. Ltd., London, U.K., 2009.

[4] C.-L. Chang and P.-Y. Lu, ‘‘The study on evaluating length of hospital stay for myomectomy,’’ Int. J. Sci. Eng. Invest., vol. 5, no. 59, pp. 157–162, 2016.

[5] M.-T. Chuang, Y.-H. Hu, C.-F. Tsai, C.-L. Lo, and W.-C. Lin, ‘‘The identification of prolonged length of stay for surgery patients,’’ in Proc. IEEE Int. Conf. Syst., Man, Cybern., Oct. 2015, pp. 3000–3003.

[6] C. Combes, F. Kadri, and S. Chaabane, ‘‘Predicting hospital length of stay using regression models: Application to emergency department,’’ in Proc. 10th Conférence Francophone de Modélisation, Optimisation et Simulation (MOSIM), 2014. [Online]. Available: <https://hal.archivesouvertes.fr/hal-01081557>

[7] E. El-Darzi, R. Abbi, C. Vasilakis, F. Gorunescu, M. Gorunescu, and P. Millard, ‘‘Length of stay-based clustering methods for patient grouping,’’ in Intelligent Patient Management (Series Studies), vol. 189. Berlin, Germany: Springer, 2009, pp. 39–56. VOLUME 9, 2021 44679 J. M. P. Gutiérrez et al.: Predicting LOS Across Hospital Departments

[8] E. Yasinski, D. Reilly, N. Duggal, B. S. Walker, E. Carpintero, S. Nag, and C. Hentz, ‘‘Understanding & predicting length of stay (LOS) using machine learning,’’ in Proc. Dexur. 15th Floor, 575 5th Avenue, New York, NY, USA, vol. 10017, 2017. [Online]. Available: https://dexur.com/a/mlresearch-los/6/