HOSPITAL ADMISSION ANALYSIS AND PREDICTION

[1] A Rajashekar, [2] C Shailaja, [3] G Shiva Kumar Reddy, [4] G Abhinav, [5] G Tharun
[1]Assistant Professor, [2345]Students
Department of CSE
Siddhartha Institute of Technology & Sciences, Narapally
alladirajashekar_cse@siddhartha.co.in, 23tq1a0517@siddhartha.co.in, 23tq1a0554@siddhartha.co.in, 23tq1a0518@siddhartha.co.in, 23tq1a0521@siddhartha.co.in,

Abstract

The Hospital Admission Analysis and Prediction system is designed to analyze hospital admission data and predict the likelihood of patient admission based on historical healthcare information. Hospitals generate large amounts of patient data related to demographics, medical conditions, symptoms, previous visits, test results, and treatment information. Analyzing this data effectively can help healthcare organizations understand admission patterns and improve resource planning. The proposed system uses historical hospital records to identify important factors associated with patient admissions. The collected data is first preprocessed by handling missing values, duplicate records, inconsistent entries, and categorical attributes. Exploratory data analysis is performed to understand admission trends based on patient characteristics, medical conditions, and other relevant factors. Data visualization techniques are used to present admission statistics and patterns through charts, graphs, and dashboards. Machine-learning algorithms such as Logistic Regression, Decision Tree, Random Forest, and Support Vector Machine can be applied to build a prediction model. The trained model analyzes relevant patient information and predicts the possible admission outcome for new cases. The performance of the prediction models can be evaluated using metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. The system can help hospitals estimate potential admission demand and support better allocation of beds, staff, and other resources. It can also assist healthcare administrators in identifying admission patterns and planning hospital operations more efficiently. The system is intended as a decision-support tool and not as a replacement for clinical judgment or medical diagnosis. Overall, the Hospital Admission Analysis and Prediction system demonstrates how data analytics and machine learning can be used to analyze healthcare admission patterns, generate predictions, and support efficient hospital resource planning and decision-making.

I. Introduction

The healthcare sector generates a large amount of information through patient visits, hospital admissions, medical records, laboratory reports, diagnoses, treatments, and other healthcare activities. Managing and analyzing this information effectively is important for improving hospital operations and resource planning. Hospitals frequently experience variations in the number of patients admitted over different days, weeks, and months. Therefore, analyzing historical admission data can help healthcare administrators understand admission patterns and prepare resources more effectively. The Hospital Admission Analysis and Prediction system is developed to analyze historical hospital data and predict possible admission outcomes based on relevant patient and hospital-related attributes.

Hospital admission decisions and patterns can be influenced by several factors, including patient age, gender, medical condition, previous hospital visits, symptoms, admission type, length of stay, and other available attributes. By studying these factors, the system can identify relationships and patterns within historical admission records. Such analysis can provide useful information about patient admission trends and help administrators understand changes in hospital demand. The system focuses on analyzing available historical data rather than replacing the professional judgment of healthcare providers.

The system begins with the collection of historical hospital admission data from a suitable dataset or hospital information source. The collected data may contain missing values, duplicate records, inconsistent information, and categorical attributes. Data preprocessing techniques are applied to clean the dataset and convert it into a suitable format for analysis. Missing values and inconsistent records are handled appropriately, while categorical attributes can be encoded for further processing. This preprocessing stage helps improve the quality and reliability of the data used by the system.

After preprocessing, Exploratory Data Analysis (EDA) is performed to understand the characteristics of hospital admission data. Statistical analysis and visualization techniques can be used to study admission frequency, patient demographics, medical conditions, admission types, and other relevant factors. Charts, graphs, and tables can provide a clear representation of admission trends and relationships between different variables. These analytical results can help identify important patterns within the historical dataset.

Machine-learning techniques can then be applied to develop a prediction model using the processed hospital data. Classification algorithms such as Logistic Regression, Decision Tree, Random Forest, and Support Vector Machine can be considered depending on the prediction objective and dataset characteristics. The models are trained using historical records and evaluated using suitable performance measures such as accuracy, precision, recall, F1-score, and ROC-AUC. The best-performing model can then be used to generate predictions for new data.

II. Literature Survey

The use of data analytics and machine learning in healthcare has increased significantly because hospitals generate large volumes of patient and operational data. Researchers have explored predictive models for hospital admission, patient outcomes, readmission, and healthcare resource utilization. These studies show that historical patient information can be analyzed to identify patterns and support healthcare planning and decision-making. Machine-learning approaches are particularly useful when healthcare datasets contain multiple demographic, clinical, and administrative variables.

Traditional hospital admission analysis mainly relies on statistical methods and historical reporting. Statistical techniques can be useful for understanding admission trends, seasonal variations, patient demographics, and disease patterns. However, conventional approaches may have limitations when dealing with large and complex

datasets. Machine-learning methods provide an opportunity to discover nonlinear relationships between variables and generate predictions from historical records.

Several studies have investigated machine-learning models for predicting hospital admissions and healthcare utilization. Classification algorithms such as Logistic Regression, Decision Trees, Random Forest, Support Vector Machines, and Gradient Boosting have been applied to patient-related datasets. These models can use factors such as age, previous healthcare utilization, medical conditions, symptoms, and other available patient information to estimate the probability of future healthcare events. The results of such studies indicate that predictive analytics can support healthcare organizations in identifying patients or periods associated with higher demand.

Research on hospital readmission prediction is also closely related to hospital admission analysis. Researchers have used electronic health records and administrative healthcare data to predict whether patients are likely to return to a hospital after discharge. Studies have explored both traditional statistical models and advanced machine-learning techniques for this purpose. The literature indicates that variables such as previous admissions, diagnoses, length of stay, comorbidities, and healthcare utilization can provide useful predictive information.

Recent research has increasingly focused on ensemble learning and deep-learning techniques for healthcare prediction. Methods such as Random Forest, Gradient Boosting, XGBoost, neural networks, and other advanced models can capture complex relationships within healthcare data. Ensemble models are particularly useful because they combine predictions from multiple models and can provide strong predictive performance. However, highly complex models may be difficult to interpret, which is an important consideration when predictions are used in healthcare environments.

Another important area of research is healthcare resource and demand forecasting. Hospitals need to plan the availability of beds, medical staff, equipment, and other resources according to expected patient demand. Researchers have therefore investigated time-series analysis and machine-learning approaches for forecasting hospital admissions and patient volumes. Accurate forecasting can help hospital administrators prepare resources in advance and reduce operational problems caused by unexpected increases in patient demand.

The literature also emphasizes the importance of data preprocessing and model evaluation in healthcare prediction systems. Healthcare datasets frequently contain missing values, inconsistent records, categorical variables, and imbalanced target classes. Therefore, preprocessing techniques such as missing-value treatment, encoding, normalization, feature selection, and class-balancing methods may be required before model training. Models can be evaluated using metrics such as accuracy, precision, recall, F1-score, ROC-AUC, and confusion matrices rather than relying only on accuracy.

III. System Analysis

The **Hospital Admission Analysis and Prediction** system is designed to analyze historical hospital admission data and predict possible admission outcomes. The

system considers important patient and admission attributes such as age, gender, medical condition, admission type, previous visits, and other available information. Historical hospital records are used to identify patterns and trends associated with patient admissions. The system first collects and organizes the available admission data for further processing. Data preprocessing is performed to handle missing values, duplicate records, inconsistent information, and categorical attributes. Exploratory Data Analysis is conducted to understand the distribution and relationships between different variables. Data visualization techniques are used to represent admission patterns through charts, graphs, and dashboards. Relevant features are selected and transformed into a suitable format for machine-learning algorithms. Different machine-learning models can be trained using historical admission records to identify useful patterns. The trained model can generate predictions for new patient or admission-related data. Model performance can be evaluated using metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. Overall, the system converts hospital admission data into meaningful analytical and predictive information to support efficient hospital planning and decision-making.

Existing System

In the existing system, hospitals generally maintain patient and admission information using hospital management systems, databases, or manual records. These systems are mainly used to register patients, maintain medical information, manage admissions, and record discharge details. Hospital administrators can generate basic reports containing information about the number of admissions and patient categories. Traditional systems generally focus on recording and retrieving historical information rather than predicting future admission patterns. When the amount of patient data increases, manually analyzing the information can become difficult and time-consuming. Identifying relationships between patient characteristics and admission patterns may require additional analytical effort. Traditional reporting systems may also provide limited visualization and predictive capabilities. Hospital staff may depend on historical observations and manual calculations for planning resources. Unexpected changes in patient admission volume can therefore be difficult to anticipate. The existing approach may not effectively utilize historical data for machine-learning-based prediction. Generating detailed analytical reports may also require considerable manual processing. Therefore, traditional systems may not provide sufficient predictive support for efficient hospital admission planning and data-driven decision-making.

Disadvantages of Existing System

- Manual analysis of hospital admission data is time-consuming.
- Large healthcare datasets are difficult to analyze efficiently.
- Limited ability to predict future hospital admission patterns.
- Difficulty in identifying hidden relationships within patient data.
- Heavy dependence on manual reports and historical observations.
- Limited visualization of admission trends and patterns.
- Greater possibility of errors during manual data analysis.
- Historical data may not be fully utilized for predictive analysis.
- Difficult to analyze multiple patient attributes simultaneously.
- Hospital resource planning can be challenging without predictions.

- Generating detailed reports may require additional manual effort.
- Limited support for data-driven administrative decision-making.

Proposed System

The proposed **Hospital Admission Analysis and Prediction** system uses data analytics and machine-learning techniques to analyze historical hospital admission records and generate predictions. The system accepts relevant patient and admission information such as age, gender, medical condition, admission type, previous visits, and other available attributes. Historical admission records are used to identify patterns associated with different admission outcomes. The collected dataset is first preprocessed by handling missing values, duplicate records, inconsistent information, and categorical attributes. Exploratory Data Analysis is performed to understand patient characteristics and identify important admission trends. Data visualization techniques are used to represent admission statistics through charts, graphs, and dashboards. Relevant features are selected and transformed into a suitable format for machine-learning models. Classification or prediction algorithms can be trained using the processed historical data. Models such as Logistic Regression, Decision Tree, Random Forest, and Support Vector Machine can be considered depending on the prediction objective. The trained models are evaluated using suitable performance metrics to select an effective model. The selected model can then be used to generate predictions for new patient or admission-related data. Overall, the proposed system provides an integrated platform for hospital admission analysis and prediction that can support better resource planning and data-driven decision-making.

Advantages of Proposed System

- Automates hospital admission data analysis.
- Reduces the time required for data processing.
- Efficiently analyzes large amounts of historical admission data.
- Helps identify important admission patterns and trends.
- Provides predictions based on historical hospital data.
- Supports hospital resource and capacity planning.
- Provides charts, graphs, and dashboards for better understanding.
- Reduces repetitive manual analysis.
- Identifies relationships between patient attributes and admission outcomes.
- Allows comparison of different machine-learning algorithms.
- Provides model performance using evaluation metrics.
- Supports data-driven hospital management decisions.
- Can be extended with advanced machine-learning techniques.

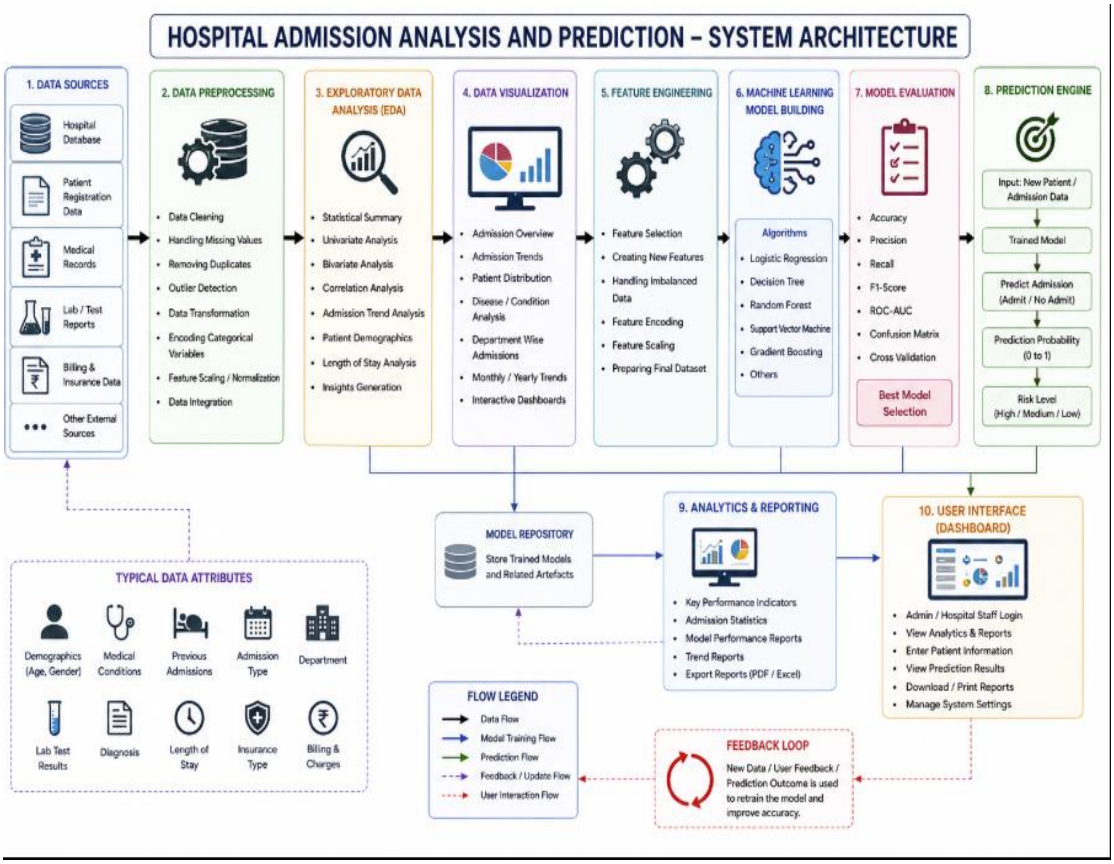
IV. Methodology

The methodology of the **Hospital Admission Analysis and Prediction** system consists of several stages, beginning with data collection and ending with prediction and result visualization. First, a historical hospital admission dataset containing patient and admission-related information is collected. The dataset is examined to understand its attributes, data types, missing values, and target variable. Data preprocessing is performed to handle missing values, duplicate records, inconsistent information, and categorical attributes. Categorical variables are converted into

numerical representations using suitable encoding techniques. Exploratory Data Analysis is performed to understand patient characteristics, admission distribution, and relationships between different variables. Data visualization techniques are used to identify important trends and patterns in hospital admissions. Relevant features are selected and prepared for training the machine-learning models. The processed dataset is divided into training and testing datasets for model development and evaluation. Different machine-learning algorithms such as Logistic Regression, Decision Tree, Random Forest, and Support Vector Machine can be trained using the prepared data. The trained models are evaluated using accuracy, precision, recall, F1-score, and ROC-AUC or other suitable performance measures. Finally, the best-performing model is selected to generate predictions, and the results are displayed through a user-friendly interface along with analytical charts and reports.

System Architecture

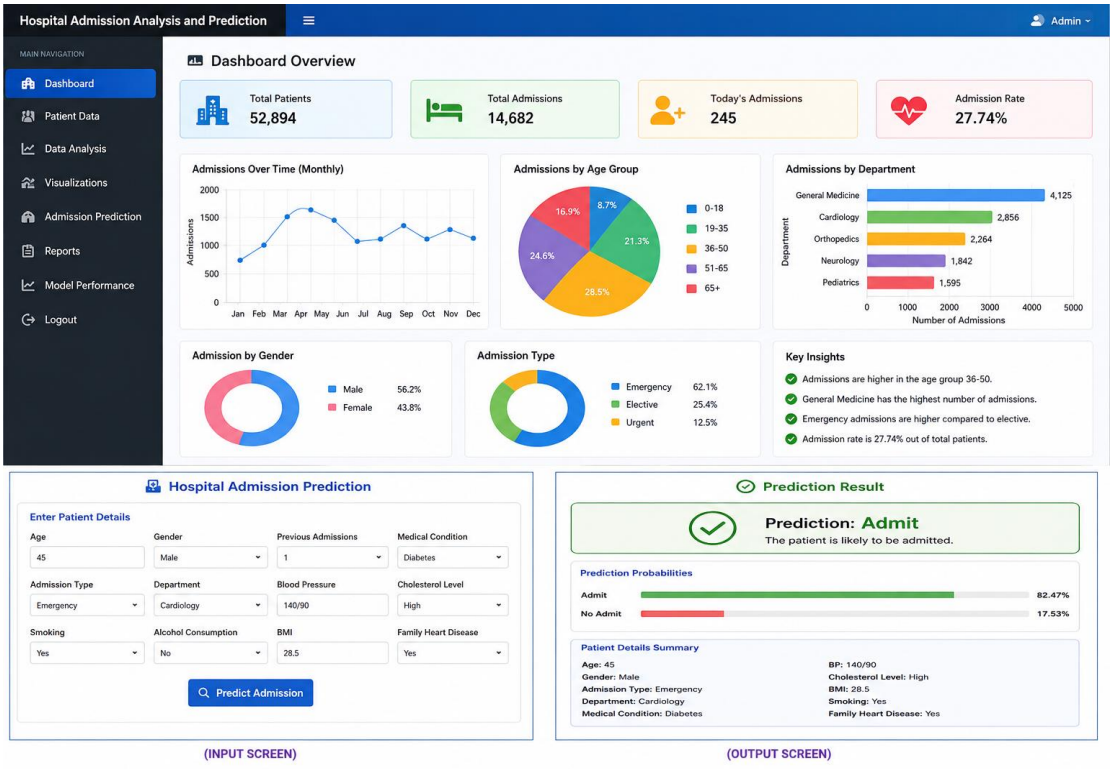
The Hospital Admission Analysis and Prediction system architecture represents the complete flow of hospital data from collection to final prediction. The system starts with the Data Collection Layer, where patient and hospital admission information is collected from historical datasets or hospital records. The collected data is transferred to the Data Preprocessing Layer, where missing values, duplicate records, inconsistent information, and categorical attributes are handled. The cleaned data is then passed to the Data Analysis Layer, where statistical analysis and exploratory techniques are used to identify admission patterns.



The Data Visualization Layer represents important admission trends and patient information using charts, graphs, and dashboards. The Feature Engineering Layer

selects the most relevant patient and admission attributes required for prediction. The processed data is then provided to the Machine Learning Layer, where suitable prediction algorithms are trained using historical admission records. The trained models are evaluated using performance metrics such as accuracy, precision, recall, and F1-score. The best-performing model is selected and connected to the Prediction Layer to predict the possible admission outcome for new data. The Reporting Layer displays prediction results, admission statistics, and analytical insights in an understandable format. Finally, the User Interface Layer allows authorized users to enter patient information and view the generated analysis and prediction results. Thus, the complete architecture follows the flow **Data Collection → Preprocessing → Analysis → Visualization → Feature Engineering → Machine Learning → Model Evaluation → Prediction → Reporting → User Interface**, providing an organized framework for hospital admission analysis and prediction.

V. Result and Output



VI. Conclusion

The HOSPITAL ADMISSION ANALYSIS AND PREDICTION system provides an effective approach for analyzing hospital admission data and predicting possible admission outcomes. The system uses historical patient and hospital records to identify important patterns and relationships related to admissions. Data preprocessing techniques help improve the quality of the dataset by handling missing values, duplicate records, inconsistent information, and categorical attributes. Exploratory Data Analysis and visualization techniques provide useful insights into patient characteristics, admission trends, medical conditions, and admission types. Machine-learning algorithms can be applied to learn patterns from historical data and generate predictions for new patient or admission-related information. Different models can be

evaluated using metrics such as accuracy, precision, recall, F1-score, and ROC-AUC to identify a suitable prediction model. The system also provides graphical dashboards and reports that make admission information easier to understand. The prediction results can assist hospital administrators in estimating admission demand and planning resources such as beds, staff, and facilities. By using historical data effectively, the system can reduce manual analysis and support faster, data-driven administrative decisions. However, the prediction should be treated as a decision-support tool and should not replace professional medical judgment or clinical diagnosis. In the future, the system can be enhanced with real-time hospital data, advanced machine-learning algorithms, time-series forecasting, and additional healthcare attributes. Overall, the project demonstrates how data analytics and machine learning can be effectively used to analyze hospital admissions, generate predictions, and support efficient hospital resource planning and decision-making.

References

[1] Kumar, R. D., Prudhviraj, G., Vijay, K., Kumar, P. S., & Plugmann, P. (2024). Exploring COVID-19 through intensive investigation with supervised machine learning algorithm. In Handbook of Artificial Intelligence and Wearables (pp. 145-158). CRC Press.

[2] Swathi, B., Vijay, K., Sushanth Babu, M., & Dinesh Kumar, R. (2024, November). Machine Learning Techniques in Cloud Based Intrusion Detection. In The International Conference on Artificial Intelligence and Smart Environment (pp. 557-564). Cham: Springer Nature Switzerland.

[3] Sv satyakrishna, shirisha rangu ,bhargavi nalacheruve.(2024) Prospective investigation on colorectal cancer with SMOTE on machine learning Algorithm

[4] Dr.G.Vishnu Murthy, BhargaviNalacheruve 1Professor, Department of computer Science & engineering, Anurag University, TS, India. 2Student, Department of computer Science & engineering, Anurag University, TS, India.

[5] V. N. S. Manaswini, K. K, C. Nigam, S. S. Ali, R. Niranjana, and Suman, "Real-Time Object Detection in Drone Surveillance Using YOLOv5," in Proc. 2025 3rd Int. Conf. IoT, Communication and Automation Technology (ICICAT), Gorakhpur, India, 2025, pp. 1–6, doi: 10.1109/ICICAT68430.2025.11414670.

[6] B. Soundarya, V. N. S. Manaswini, M. Ayyakrishnan, R. D. Kumar, "Contextual Analysis of Big Data Analytics in Intelligent Transportation Frameworks," in Intersection of Artificial Intelligence, Data Science, and Cutting-Edge Technologies: From Concepts to Applications in Smart Environment, Lecture Notes in Networks and Systems, vol. 1353, Cham: Springer, 2025, doi: 10.1007/978-3-031-88304-0_79.

[7] R. D. Kumar, V. N. S.Manaswini, "Applications of blockchain in smart cities: detecting fake documents from land records using blockchain technology," in Blockchain for Smart Cities, Elsevier, 2021, pp. 105–117, doi: 10.1016/B978-0-12-824446-3.00017-X.

[8] Tejavath Veeramma, Badarla Anil, Guguloth Ravinder, "An advanced movie recommender using collaborative filtering and sentiment analysis," International Research Journal of Modernization in Engineering Technology and Science, vol. 7, no. 7, July 2025, doi: 10.56726/IRJMETS81618.

[9] Ravi Kumar Banoth, Ramana Murthy B V, "Automatic crop recommendation system using LightGBM and decision tree machine learning models," Journal of Machine and Computing, vol. 5, no. 1, pp. 343, Jan. 2025, doi: 10.53759/7669/jmc202505026.

[10] Ravi Kumar Banoth, Dr. B.V. Ramana Murthy, "Smart agriculture through IoT and machine learning for analyzing carbon footprints," in Proc. Int. Conf. Computer Science and Communication Engineering (ICCSCE), Apr. 2025.
[11] Ravi Kumar Banoth, B. V. Ramana Murthy, "Soil image classification using transfer learning approach: MobileNetV2 with CNN," SN Computer Science, vol. 5, art. no. 199, 2024, doi: 10.1007/s42979-023-02500-x.