PREDICTIVE ANALYSIS ON AVAILABILITY OF DOCTORS AND MEDICINES IN GOVERNMENT HOSPITALS

A PROJECT REPORT

Submitted by,

Ms. SAHANA REDDY R -20211CSD0192
Ms. DEEPIKA R -20211CSD0064
Ms. LISHA S -20211CSD0063

Under the guidance of,

DR. SRABANA PRAMANIK

in partial fulfillment for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND ENGINEERING (DATA SCIENCE)

At

GAIN MORE KNOWLEDGE
REACH GREATER HEIGHTS

PRESIDENCY UNIVERSITY

BENGALURU

JANUARY 2025

CERTIFICATE

This is to certify that the Project report **"Predictive Analysis on Availability of doctors and medicines"** being submitted by "SAHANA REDDY R, DEEPIKA R, LISHA S" bearing roll number(s)"20211CSD00192, 20211CSD0064, 20211CSD0063" in partial fulfillment of the requirement for the award of the degree of Bachelor of Technology in Computer Science and Engineering is a Bonafide work carried out under my supervision.

Dr. SRABANA PRAMANIK
Assistant Professor
Presidency School of CSE
Presidency University

Dr. SAIRA BANU ATHAM
Professor & HoD
Presidency School of CSE
Presidency University

Dr. L. SHAKKEERA
Associate Dean School of CSE
Presidency University

Dr. MYDHILI NAIR
Associate Dean School of CSE
Presidency University

Dr. SAMEERUDDIN KHAN
Pro-Vc School of Engineering Dean – Presidency School of CSE&IS
Presidency University

DECLARATION

We hereby declare that the work, which is being presented in the project report entitled "PREDICTIVE ANALYSIS ON AVAIABILITY OF DOCTORS AND MEDICINES IN GOVERNMENT" in partial fulfillment for the award of Degree of **Bachelor of Technology** in **Computer Science and Engineering (DATA SCIENCE)**, is a record of our own investigations carried under the guidance of **Dr. SRABANA PRAMANIK, ASSISTANT PROFFESSOR , School of Computer Science Engineering and Information Science , Presidency University, Bengaluru.**

We have not submitted the matter presented in this report anywhere for the award of any other Degree.

SAHANA REDDY R -20211CSD0192
DEEPIKA R-20211CSD0064
LISHA S-20211CSD0063

ABSTRACT

In the modern era where the world is driven by modern technology in almost every field, the Healthcare field is not far behind and is rapidly improving, marked by integration of various algorithm and technology. This abstract will discuss how the predictive analysis will benefit the healthcare industry and how the traditional system can be transformed into an efficient system. This paper will also shed light on how the healthcare system which primarily follows the clinical approach for diagnosis can be digitalized using the appropriate analysis algorithm to ensure the information, availability of doctors and medicines, patient database and other activities. In the system that we developed, users can register with their details, and they will be stored in admin's database. Users will be able to view various information regarding hospitals and doctors to their best interest.

The enhanced utilization of resources, along with measures to minimize fraud and abuse, are contributing factors for financial performance and administrative outcome. A proper strategic and efficient healthcare information system of predictive nature amalgamated with tons of useful features running on cutting-edge technology affordable by all classes of society will prove to be a milestone in the public domain. Overall, the contribution of predictive analysis towards the healthcare system is significant. Government hospitals provide medicines for the treatment to the patients based on the diagnosis. Generally, government hospitals store all the patient's historical data and current data in cloud. In our system user can register with there details, which are stored in the admin's database. This system allows the user to view the hospital location using predictive algorithm and details about the hospital such as doctors, Medicines, specialists' availability and helps the patient to get details about the government hospitals. High utilization of resources and reduced fraud and abuse optimize financial and administrative performance. Supply chain and human capital management optimize the financial and administrative performance. Government hospitals are the main providers of health care to a large population, and health care accessibility is an important aspect of public welfare. The issues that are critical and require a solution are the adequate supply of essential medicines and the distribution of doctors in government hospitals, and the framework is a predictive analysis based on this study. The framework analyzes historical data, applies machine learning algorithms, and employs statistical models to uncover patterns and trends that impact availability and resource allocation Future demand and potential shortages are predicted through analyses of critical factors such as patient inflow, disease seasonality, inventory logistics, and staff scheduling.

ACKNOWLEDGEMENT

First, we indebted to the **GOD ALMIGHTY** for giving me an opportunity to excel in our efforts to complete this project on time.

We express our sincere thanks to our respected dean **Dr. Md. Sameeruddin Khan**, Pro-VC, School of Engineering and Dean, School of Computer Science Engineering & Information Science, Presidency University for getting us permission to undergo the project.

We express our heartfelt gratitude to our beloved Associate Deans **Dr. Shakkeera L and Dr. Mydhili Nair,** School of Computer Science Engineering & Information Science, Presidency University, and Dr. “SAIRA BANU ATHAM”, Head of the Department, School of Computer Science Engineering & Information Science, Presidency University, for rendering timely help in completing this project successfully.

We are greatly indebted to our guide **Dr. SRABANA PRAMANIK, ASSISTANT PROFESSOR** and Reviewer **Dr. RADHIKA SHARMA, ASSISTANT PROFESSOR**, School of Computer Science Engineering & Information Science, Presidency University for his/her inspirational guidance, and valuable suggestions and for providing us a chance to express our technical capabilities in every respect for the completion of the project work.

We would like to convey our gratitude and heartfelt thanks to the PIP2001 Capstone Project Coordinators **Dr. Sampath A K, Dr. Abdul Khadar A and Mr. Md Zia Ur Rahman,** department Project Coordinators **Mrs. Manjula S G** and Github coordinator **Mr. Muthuraj.**

We thank our family and friends for the strong support and inspiration they have provided us in bringing out this project.

SAHANA REDDY R -20211CSD0192

DEEPIKA R-20211CSD0064

LISHA S-20211CSD0063

LIST OF TABLES

LIST OF FIGURES

LIST OF CONTENTS

CHAPTER 1

INTRODUCTION

The usage of technical and scientific tools in the healthcare system has exponentially increased in the last decade. This revolutionized the entire system by integrating predictive analytics and the internet, linking our health with personal information. This integration is applied in various aspects of our day-to-day life, both superficially and in-depth.

1.1. Healthcare Management

Managing doctors and medicines during peak times in congested hospitals is a significant challenge. With hospitals facing high patient inflows, managing the availability of specialists and medicines becomes difficult. The application of predictive analytics can analyze a patient's past and present data to generate short analyses on medicines and the required specialists, streamlining hospital operations and improving management efficiency.

1.2. Predictive Analysis for Efficient Hospital Management

This system can efficiently manage hospital workflows by predicting doctor availability, patient attendance, and even when patients might miss appointments. By integrating algorithms with historical and real-time data, it allows the hospital administration to reduce unexpected absences and manage resources proactively. This predictive capability ensures timely appointments and efficient doctor and patient allocation.

1.3. Benefits of Predictive Analysis in Healthcare

Predictive analysis in healthcare allows for the early prediction of disease patterns, enabling preventive measures. By tracking patients' historical records, the system can predict vulnerabilities to certain diseases and advise patients accordingly. This reduces the chances of missed appointments and improves patient engagement with their healthcare providers.

1.4. Machine Learning Models for Predictive Analysis

The system uses advanced machine learning models, including Random Forest, Decision Tree, and Convolutional Neural Networks (CNN), to forecast the availability of medicines and doctors. These models help in predicting medicine demand and doctor availability, which in turn assists in making informed decisions regarding resource allocation and management.

1.5. Impact on Patient Care and Health Outcomes

The use of predictive analysis significantly improves patient care by ensuring that doctors are available when needed and that medicines are in stock. It also reduces the likelihood of missed appointments, which can lead to a negative reputation for the hospital. Additionally, the system ensures that patients are informed about their health and appointments, improving patient satisfaction and care outcomes.

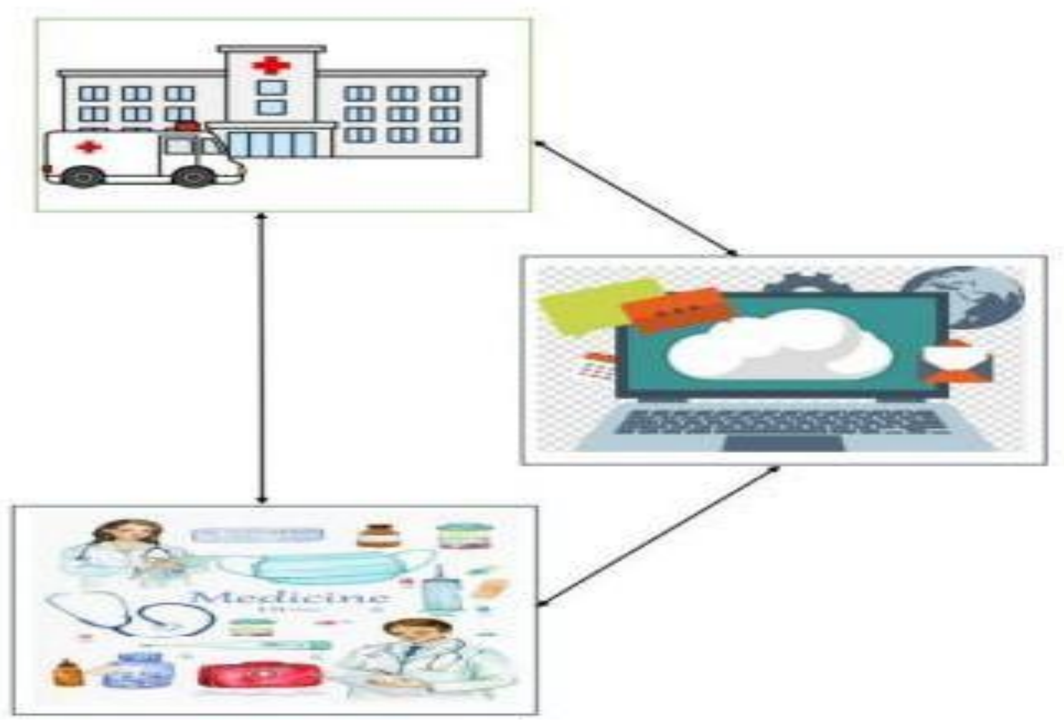


Figure 1.1 Healthcare Information Technology

1.6. Challenges in Traditional Inventory Management

Traditional inventory management systems in hospitals are often reactive, relying on historical consumption data and manual interventions. These systems are prone to errors like overstocking or understocking, which can lead to wasted resources or failure to meet patient needs. Managing hospital drug supplies has become a major challenge due to the variability in demand caused by seasonal diseases and patient demographics.

1.7. Role of Predictive Analytics in Inventory Management

By applying machine learning models to predict future medicine demand, hospitals can forecast trends more accurately and avoid supply shortages or overstocking. Predictive analytics enables hospitals to identify hidden patterns in data and take proactive measures to manage their inventory more effectively.

1.8. Key Features of the Predictive Analysis System

Key features of the system include:

1. **Machine Learning Models**: Using sophisticated algorithms to analyze data and predict medicine demand and doctor availability.
2. **Data Integration**: The system integrates various data sources for more accurate predictions.
3. **Proactive Resource Management**: Ensures that essential medicines are available, and healthcare professionals are optimally distributed.
4. **Interactive Dashboard**: A user-friendly interface for administrators to view predictive insights and manage resources.
5. **Community Impact Assessment**: The system evaluates its impact on healthcare accessibility and equity, especially in underserved areas.

CHAPTER 2

LITERATURE SURVEY

Predictive analytics has emerged as a crucial tool in healthcare, addressing issues such as medicine shortages and optimizing resource allocation. By using ML models, healthcare systems can predict patient inflow and disease trends, which helps hospitals maintain adequate stocks of critical medicines. Regression algorithms, as used in this study, are the backbone of predictive systems, correlating historical and real-time patient data to identify patterns and predict future requirements. Big data technologies, like Hive, when integrated with Python libraries like NumPy, Pandas, Scikit-learn, and Matplotlib, enhance the efficiency and accuracy of data processing and analysis. These tools facilitate preprocessing, modeling, and visualization, offering insights that support decision-making. Previous research has demonstrated the efficacy of predictive analytics in various domains, including anticipating infection risks, monitoring communicable diseases, and optimizing hospital readmissions. Studies also highlight the significance of EHRs in holding and analyzing big data to identify relationships in health care data.

The incorporation of predictive models with big data analytics has been quite effective in bridging the gaps between resource availability and patient needs. For example, the analysis of real-time data helped to reduce the shortage of medicines during peak periods of disease, thus enabling patients to receive timely treatment and improved outcomes. Moreover, predictive tools help the government and health agencies make data-driven decisions to enhance service delivery and public health infrastructure. This literature underlines the importance of machine learning and big data in transforming healthcare into a proactive and efficient system that can predict and address problems before they become unmanageable. By building on these foundations, the proposed system of this study aims to revolutionize medicine management in government hospitals, ensuring availability during critical times and ultimately improving the quality of care.

Predictive analytics is changing healthcare as it helps assess vast datasets to uncover patterns, predict outcomes, and optimize decision-making processes. Sophisticated techniques, including regression models, data mining, and machine learning, provide insights into patient management and resource allocation. Factors such as age, prematurity, and hemoglobin levels have been shown to be major predictors of ICU stays for pediatric cardiac surgery patients; hence, the necessity for data-driven preoperative planning. Predictive models improve clinical efficiency and save costs by identifying risks and reducing postoperative complications. Fraud detection, marketing, and preventive medicine are other applications of predictive analytics that demonstrate its utility in improving healthcare outcomes. With healthcare organizations digitizing their data, predictive tools allow for customized treatments, better resource allocation, and informed policy decisions that shift the focus toward precision medicine and better patient care

Shaikh Karnool Afsa et al. [10] in their article explores predictive analytics has emerged as a crucial tool in healthcare, addressing issues such as medicine shortages and optimizing resource allocation. By using ML models, healthcare systems can predict patient inflow and disease trends, which helps hospitals maintain adequate stocks of critical medicines. Regression algorithms, as used in this study, are the backbone of predictive systems, correlating historical and real-time patient data to identify patterns and predict future requirements. Big data technologies, like Hive, when integrated with Python libraries like NumPy, Pandas, Scikit-learn, and Matplotlib, enhance the efficiency and accuracy of data processing and analysis. These tools facilitate preprocessing, modeling, and visualization, offering insights that support decision-making. Previous research has demonstrated the efficacy of predictive analytics in various domains, including anticipating infection risks, monitoring communicable diseases, and optimizing hospital readmissions. Studies also highlight the significance of EHRs in holding and analyzing big data to identify relationships in health care data. The incorporation of predictive models with big data analytics has been quite effective in bridging the gaps between resource availability and patient needs. For example, the analysis of real-time data helped to reduce the shortage of medicines during peak periods of disease, thus enabling patients to receive timely treatment and improved outcomes. Moreover, predictive tools help the government and health agencies make data-driven decisions to enhance service delivery and public health infrastructure. This literature underlines the importance of machine learning and big data in transforming healthcare into a proactive and efficient system that can predict and address problems before they become unmanageable. By building on these foundations, the proposed system of this study aims to revolutionize medicine management in government hospitals, ensuring availability during critical times and ultimately improving the quality of care.

M.D Boomija et al [2] in their article explores that integration of predictive analytics has helped improve the operational efficiency of government hospitals in large measures. Predictive models with the help of historical as well as current patient data help to estimate medicine requirements and optimize the availability of doctors and specialists. The analytics of big data will identify variables influencing resource allocation and develop strategies for improved service delivery. Advanced tools such as R, HTML, and predictive algorithms like Random Forest enable the development of robust models for decision-making. These innovations ensure timely patient care, reduce human effort, and streamline hospital operations. Studies highlight the importance of data analytics in improving healthcare outcomes, identifying patient needs, and enhancing the efficiency of public health systems. The adoption of predictive analytics also helped in the better resources management, fraud reduction as well as improved patient satisfaction owing to timely access to medicines and specialists.

Manish Kumar et al [3] in the article explores that the paper delves on the application of predictive analytics into healthcare, which would help develop better quality care at relatively lower costs. It takes into account data mining techniques - classification, association, and clustering - to address all healthcare issues like risk estimation, patient tracking, fraud detection, and so many more. The study encompasses tools such as the Charlson Comorbidity Index estimating health risks and offers a comparative review of existing predictive tools in use. The paper emphasizes transitioning healthcare systems from reactive to proactive approaches, with predictive analytics reducing readmissions and enabling cost-efficient, high-quality care. Challenges include privacy concerns, data integration, and standardization.

Smitha Jhajharia et al [4] in the article explores predictive analytics is changing healthcare as it helps assess vast datasets to uncover patterns, predict outcomes, and optimize decision-making processes. Sophisticated techniques, including regression models, data mining, and machine learning, provide insights into patient management and resource allocation. Factors such as age, prematurity, and hemoglobin levels have been shown to be major predictors of ICU stays for pediatric cardiac surgery patients; hence, the necessity for data-driven preoperative planning. Predictive models improve clinical efficiency and save costs by identifying risks and reducing postoperative complications. Fraud detection, marketing, and preventive medicine are other applications of predictive analytics that demonstrate its utility in improving healthcare outcomes. With healthcare organizations digitizing their data, predictive tools allow for customized treatments, better resource allocation, and informed policy decisions that shift the focus toward precision medicine and better patient care

Damien S.E Brokharst et al [5] in the article explores the role that predictive analytics plays in improving the agility and value of a patient within a hospital system. It discusses the reactionary ability of current hospital capabilities, focusing on the need for an active, predictive approach in healthcare. Predictive analytics enables hospitals to sense patient needs before problems arise, preventing negative trends and better patient outcomes. The research proposal calls for a multi-stakeholder perspective and a comprehensive conceptual framework that integrates biomedical and health service needs. Using predictive analytics, hospitals can enhance evidence-based medical practices, manage capacity, and streamline healthcare pathways. The paper also outlines the potential for an innovation ecosystem that facilitates collaborative data exchange among stakeholders to foster improved healthcare delivery. Future research directions include examining predictive analytics in diverse healthcare settings, exploring related analytics types, and optimizing resource allocation to maximize patient agility and value.

Sharique Ahmad et al [6] in the article explores the Advanced analytics, predictive and big data analytics, transform the healthcare landscape by enabling data-driven decision-making. Predictive analytics uses historical medical data, statistical models, and machine learning to predict future outcomes, thus being a proactive approach to health care management. Key applications include disease prediction, personalized treatment plans, resource optimization, and fraud detection. For instance, predictive models have been successfully deployed in identifying patients who are at a high risk of readmission or sepsis. This saves costs and ensures better outcomes. Big data analytics, in turn, applies the "4 Vs," which include volume, velocity, variety, and veracity, in extracting actionable insights from a diverse set of datasets like electronic health records and social media. The technologies Hadoop and machine learning provide real-time analytics, hence improving care quality and operational efficiency. Despite the promise, the integration of these technologies is challenged by data privacy, standardization, interoperability, and algorithmic bias, among other ethical issues. Interdisciplinary approaches are suggested to address these problems, including compliance with regulatory frameworks and improving the interpretability of black-box models. The recent advancements that include genomic data integration and telehealth applications highlight the growing role of predictive analytics in precision medicine and remote monitoring. However, there are still challenges such as data fragmentation and organizational resistance to the widespread adoption of this technology. Success stories, such as resource optimization using predictive analytics during the COVID-19 pandemic, highlight the potential for revolutionizing patient care. Future research is focused on improving model transparency, addressing data quality issues, and integrating emerging technologies like IoT and AI to promote holistic, personalized healthcare systems. Predictive and big data analytics are enormous in their promise to innovate, improve patient outcomes, and streamline healthcare delivery with further progress and ethical vigilance.

Mohammed Badawy et al [7] in the article explores that the paper is a thorough review on the integration of ML and DL techniques into predictive healthcare analytics. The author draws attention to the fact that AI is transforming the realm of medical diagnostics with respect to early disease detection and the design of personalized treatment planning. Different types of ML models including linear regression, decision trees, and random forests, besides DL architectures like CNNs and LSTM, are discussed to establish their effectiveness in healthcare prediction. Some of the challenges in using such models include dealing with large, heterogeneous datasets, achieving accuracy, and eliminating biases in the prediction algorithms. The study also delves into supervised, unsupervised, and reinforcement learning models, which can be applied to prognosis, diagnosis, therapy optimization, and enhancement of clinical workflow. Future work is in the refinement of these models to improve scalability, reduce computational complexity, and ensure ethical data governance in healthcare applications.

Wullinarllur Raghupathi et al [8] in the article explores that this paper will outline the transformative potential of big data analytics in healthcare through using vast, diverse datasets to glean insights that improve outcomes and reduce costs. The paper defines key concepts such as the "4 Vs" of volume, velocity, variety, and veracity, and describes frameworks like Hadoop for managing and analyzing healthcare data. Examples of successful applications include disease surveillance, fraud detection, and personalized medicine. Despite its promise, challenges remain in data standardization, privacy, and skill gaps. The paper concludes by advocating for advanced platforms, tools, and policies to realize the full potential of big data in healthcare.

Yichung Wang et al [9] in the article explores that the document is a discussion of the potential strategic applications of big data in the healthcare industry. While previous studies have mostly concentrated on the technological aspects of big data, this paper is based on its strategic implications to bridge the gap between technical capabilities and healthcare management needs.
The authors describe the architecture and functionalities of big data that enable one to process vast quantities of disparate data with various platforms like Hadoop and NoSQL systems. All these are meant to facilitate integrating, transforming, and storing data and help health institutions analyze both structured and unstructured data effectively.

Anu priya et al [10] in the article explores that the document is a discussion of the potential strategic applications of big data in the healthcare industry. While previous studies have mostly concentrated on the technological aspects of big data, this paper is based on its strategic implications to bridge the gap between technical capabilities and healthcare management needs.
The authors describe the architecture and functionalities of big data that enable one to process vast quantities of disparate data with various platforms like Hadoop and NoSQL systems. All these are meant to facilitate integrating, transforming, and storing data and help health institutions analyze both structured and unstructured data effectively. Based on such research, a study can point out important.

CHAPTER 3

RESEARCH GAPS OF EXISTING METHOD

Predictive models and systems of healthcare resource management have made rapid strides. But still, the gaps in such research are countless that limit the efficiency, scalability, and their real-world application. Some of the most noteworthy gaps are:

3.1. Quality and Availability of Data

- **[1]. Low access to overall data:** Most of the systems depend on either incomplete or artificially generated datasets that provide less precise predictions. There is a scarcity of data, especially regarding the inflow of patients, trends for medicine usage, and schedules for doctors.
- **[2]. Data Imbalance:** Some features, like rare diseases or specific medicines, are underrepresented, and thus the predictions are biased.
- **[3]. RealTime Data Integration:** Current approaches hardly use real time data from hospital information systems, IoT devices, or patient monitoring systems.

3.2. Scalability and Adaptability

- **[1]. Limited Feature Set:** Current models mainly focus on a limited set of features like stock levels or doctor availability. They do not consider auxiliary factors like patient satisfaction, disease prevalence, or seasonal variations.
- **[2]. Static Models:** Many methods rely on static training, meaning models do not adapt to evolving trends or new data without manual retraining.

3.3. Generalization Across Environments

- **[1]. Geographic and Institutional Bias:** Models trained on data from specific hospitals or regions often fail to generalize to other settings with different workflows or patient demographics.
- **[2]. Specialization Gaps:** Methods that work well for general hospitals may not be effective for specialized clinics or rural healthcare centers.

3.4. Lack of Advanced Machine Learning Techniques

[1]. **Underutilization of Deep Learning**: Often, traditional models such as Random Forests and logistic regression are used, but deep learning, which can enhance accuracy for large data, is less explored.

[2]. **Explainability and Interpretability:** Most of the models act as black boxes, providing predictions without explanations, thereby reducing trust and usability among the healthcare professionals.

[3]. **Integration of Multi-Modal Data:** Current approaches often fail to incorporate different types of data, such as imaging, text (for example, prescriptions), and timeseries data (for example, patient vitals).

3.5. Human-Centric Considerations

[1]. **User Interface Limitations:** Healthcare professionals are unlikely to use intuitive interfaces. Language barriers and accessibility limitations are not fully addressed.

[2]. **Behavioral and Cultural Factors:** Models do not capture human factors such as doctor preference, patient adherence, or cultural views on drugs and treatment.

3.6. Evaluation and Validation Challenges

[1]. **Limited Benchmarking:** Lack of standardized datasets and evaluation criteria, making it challenging to compare models across studies.

[2]. **Real-World Testing:** Most methods are validated on historical datasets but not deployed in live healthcare settings for feedback and refinement.

[3]. **Overfitting to Training Data:** Systems perform well on training data but fail to maintain accuracy when exposed to new, unseen cases.

3.7. Integration with Existing Systems

[1]. **Interoperability Issues:** Current solutions often fail to seamlessly integrate with existing Electronic Health Record (EHR) systems and hospital workflows.

[2]. **Fragmented Tools:** Healthcare systems use different tools for different functions, such as scheduling and inventory management, which leads to redundancy and inefficiency.

3.8. Ethical and Regulatory Challenges

[1]. **Bias in Predictions:** Models may inadvertently perpetuate biases related to gender, ethnicity, or socioeconomic status.

[2]. **Data Privacy Concerns:** Patient data is sensitive, and current methods often lack robust privacy-preserving-mechanisms.

[3]. **Regulatory Compliance:** Adherence to standards like HIPAA or GDPR is inconsistent, especially if the system is being implemented across various countries.

CHAPTER-4

OBJECTIVES

4.1. Healthcare Resource Availability

Goal: To help users find whether doctors and drugs are available in a hospital or clinic setting with predictions.

How:

Doctor availability is predicted based on whether they are associated with a hospital, specialty, and other categorical variables.

For drugs, predictions take stock amount and pricing information into consideration to make predictions about availability.

4.2. Building Machine Learning Models

Goal: Utilize machine learning models for predicting availability to maximize efficiency and precision.

Steps:

The `RandomForestClassifier` is applied in binary classification for the prediction of doctor availability.

A `RandomForestRegressor` is used for regression-based tasks to predict the availability of drugs.

These models are trained on historical and randomly assigned data to simulate real-world scenarios.

4.3. Dataset Preparation

Goal: To prepare data in such a way that it becomes suitable for machine learning tasks.

Implementation Steps:

Loads data from an Excel file (`Book 4.xlsx`).

Encode categorical columns with `LabelEncoder` for all the columns like `Doctors, Hospitals, Specialist, Medicines`.

Created target columns (`DoctorAvailable` and `MedicineAvailable`).

`DoctorAvailable`: binary data randomly assigned for training.

`MedicineAvailable` is obtained by considering the stock quantity and a boolean value specifying the availability.

4.4. Doctor Availability Prediction

Goal: Decide whether a doctor is available or not based on their hospital affiliation, the specialist they are, and other related factors.

Implementation Details:

Categorical features: `Doctors`, `Hospitals`, `Specialist`. These features are encoded and used for input of the `RandomForestClassifier`. It gives the output as a probability of availability (1 for available, 0 for unavailable). Output is given in an easy to understand format where it shows the name of the doctor, the hospital, specialty, and whether available or not.

4.5. Medicine Availability Prediction

Goal: To predict the availability of medicines based on stock quantity and price.

Implementation Details:

Features (`Medicines`, `Stock Quantity`, `Price`) are encoded and fed into `RandomForestRegressor`.

This model predicts the number, where a value of more than `0.5` is set to `Yes` for `Availability`. The actual predictions are outputted, showing what medicine, their stock status, price, and availability.

4.6. User Interface

Goal: Developing an easily understandable command line utility for end-users.

Implementation:

A `main_menu` function provides three choices:

Check availability of the doctor.

Check medicine's availability.

Quit

Users may input their selection to retrieve related predictions or exit the session.

4.7. Evaluation Metrics

Goal: Evaluate how well the models performed.

Implementation:

For the doctor availability model:

Accuracy score; accuracy score calculated as the proportion of correct predictions.

A classification report is also produced to indicate the precision, recall, F1score and support for every class.

For the medicine availability model:

Mean Absolute Error (MAE) is defined as the mean error in its predictions.

4.8. Hyperparameter Tuning

Goal: To optimize machine learning models for better accuracy and efficiency.

Implementation:

A `GridSearchCV` method is used to perform exhaustive search over specified hyperparameter values.

The model parameters that were tuned are the number of estimators, maximum depth, and minimum samples required for splitting nodes.

4.9. Randomization and Simulation

Goal: This is in place to mimic actual scenarios, which cannot be performed since realtime availability data cannot be known beforehand.

Usage:

Availability of the doctors is set to random as boolean values (`True` or `False`).

Availability of medicines is calculated based on stock levels where greater than zero means availability

4.10. Data Driven Insights

Goal: To empower decision makers in health care.

Application:

Administers could use the predictions to utilize resources appropriately, have medicines restocked, or manage doctor's schedules.

Intime predictions provide timely decisions in critical healthcare conditions.

4.11. Scalability and Flexibility

Goal: Design the system for adaptability and future expansions.

Implementation:

Additional features or categories (for example, nurse availability, patient waiting time) can be included in the model.

It is designed modularly such that improvements don't affect major changes in previously functional code

4.12. Data Security and Local Processing

Goal: Ensure local and secure processing of sensitive health information.

Implementation:

The code executes locally on locally loaded data from an Excel file.

Predictions and data transformations are done in memory, so confidentiality is ensured.

4.13. Ease of Deployment

Goal: To build a lightweight solution deployable in various contexts.

Implementation:

The script is self-contained and has minimal setup requirements (Python environment with the required libraries).

It can be integrated into larger systems or run independently on local machines.

4.14. Error Handling and Validation

Goal: To enhance the user experience through error handling and validation of input.

Implementation:

Menu input is validated to accept only specified choices.

Data preprocessing ensures correct encoding of all the features before performing any operations related to model training and prediction.

4.15. Applications in Real World

Goal: Demonstrate the utility of predictive modeling in real world healthcare scenarios.

Use Cases:

This system can be used by hospitals in order to optimize resource allocation.

Pharmacies can predict the availability of stocks and adjust the inventories accordingly.

Patients can check for the availability of doctors and medicines in stock before visiting healthcare facilities.

CHAPTER-5

PROPOSED METHODOLOGY

Methodology is sequential steps, with each step performing specific tasks on the system. The following is the methodology applied in this case:

5.1 Data Collection and Preprocessing

5.1.1 Data Source

Data is uploaded from an Excel file, which is Book 4.xlsx that contains historical information about healthcare resources, including:

Doctors: Names, affiliated hospitals, and specialties.

Medicines: Names, stock quantities, and prices.

5.1.2 Data Preprocessing

Categorical Encoding: Columns like Doctors, Hospitals, Specialist, and Medicines use Label-Encoder in order to transform text data to numerical format.

Feature Engineering:

Doctors: A binary label DoctorAvailable is created

Medicines: The stock level for medicines will get a binary label Medicine Available.

Data Splitting: The train set contains 80% and the test set contains 20% of the dataset.

5.1.3 Challenges

Handling categorical data with encoding.

Meaningful target labels for machine learning tasks.

5.2 Model Development

The system uses two machine learning models: a classification model for doctor availability and a regression model for medicine availability.

5.2.1 Doctor Availability Prediction

Model Used: RandomForestClassifier.

Features: Doctors, Hospitals, and Specialist.

Target Variable: DoctorAvailable.

Training Process: The Random Forest algorithm is trained on encoded categorical features to predict whether a doctor is available.

Output: A binary prediction indicating availability (1 for available, 0 for unavailable).

5.2.2 Medicine Availability Forecasting

Model: RandomForestRegressor.

Features: Drugs, Quantity in Stock, and Price.

Target: MedicineAvailable.

Training: The Random Forest algorithm predicts the availability of drugs using the stock and price.

Output: Continuous prediction that is interpreted as being available if greater than 0.5.

5.2.3 Model Validation

Accuracy: Proportion of correct predictions.

Precision, Recall: Proportion of correct predictions for each class.

Mean Absolute Error (MAE): Average absolute prediction error.

5.3 Model Optimization

To improve the models, hyperparameter tuning is done.

5.3.1 Tuning Parameters

RandomForestClassifier:

Number of trees (n_estimators).

Maximum depth of trees (max_depth).

Minimum samples required to split (min_samples_split).

RandomForestRegressor: Similar parameters as the classifier model, but more on regression specific optimization.

5.3.2 GridSearchCV

GridSearchCV is used to test combinations of hyperparameters and determine the best combination.

5.3.3 Optimization Benefits

Improving the models to exhibit enhanced accuracy and reliability.

Generalization of the system on unseen data.

5.4 System Design

The system follows a modular structure. This is to make sure the system is simple yet scalable.

5.4.1 Modules

Data Process Module: Provide function for loading the data, encoding, and preparation to the machine learning algorithm.

Model Trainer Module: Classification and regression models training.

User Interface Module: CLI to interact with the system.

Evaluation Module: Provides necessary metrics to assess the performance of the models.

5.4.2 Workflow

Load and preprocess.

Models are trained and tested.

Prediction is made using user input.

Results are displayed to the user.

5.5 User Interface

The user interacts with the system through a simple commandline interface (CLI).

5.5.1 Features

Option 1: Doctor Availability

Input the doctor's name, hospital, and specialty

Option 2: Medicine Availability

Input the medicine name, stock quantity, and price

Option 3: Exit

5.6 Results Interpretation

The system presents results in an easy-to-understand manner:

Availability of Doctors: Shows if each doctor is available according to the prediction.

Availability of Medicines: Shows the medicines available along with stock and price availability.

5.7 Scalability and Future Enhancements

The approach has been designed to scale:

Integration of Realtime Data Sources: The system can integrate into hospital databases or IoT devices to fetch realtime data.

Additional Features:

[1]. Predictions of nurse availability.

[2]. Live inventory management for hospital supplies.

[3]. User Interface Enhancements: Change the CLI to graphical or webbased, making it friendlier and much easier to use.

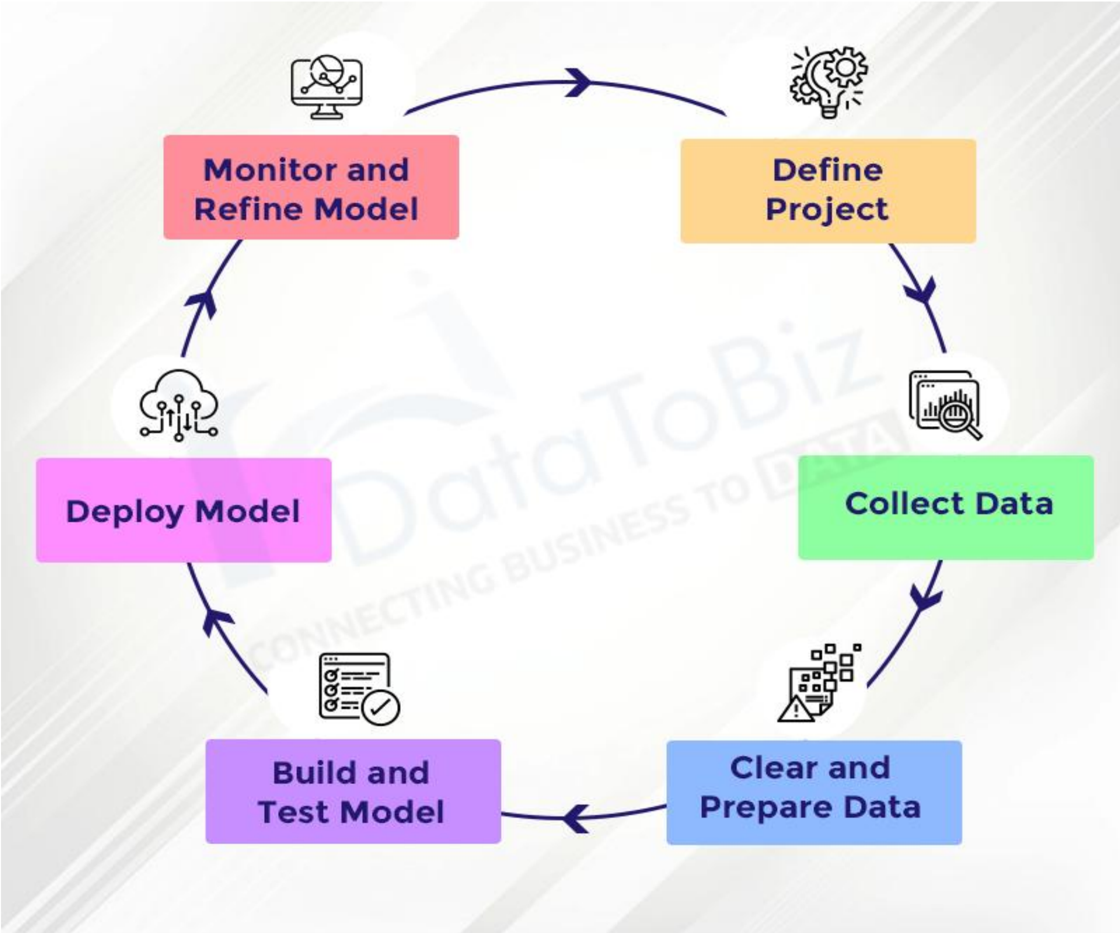


Figure 5.1. Predictive Analysis Process

CHAPTER 6

SYSTEM DESIGN & IMPLEMENTATION

6.1. System Design:

I. Essential Components

Data Source

Input: Book 4.xlsx that includes doctors, hospitals, specializations, medicines, stock quantities, and prices.

Output: Preprocessed data in a structured format, ready to process, train, and make predictions.

II. Preprocessing Module

LabelEncoder: Encoding categorical variables into numerical forms

Feature preparation and target variable preparation for the machine learning models.

III. Machine Learning Models

Doctor Availability Prediction:

Model: RandomForestClassifier

Features: Doctor's name, hospital, specialization.

Target: Binary availability status.

Medicine Availability Prediction:

Model: RandomForestRegressor

Features: Medicine name, stock quantity, price.

Target: Numerical availability prediction.

IV. Evaluation Metrics

For classification: Accuracy, precision, recall, F1-score.

For regression: Mean Absolute Error (MAE), R-squared (R^2) score.

User Interface (CLI)

Provides options for users to check the availability of doctors and medicines.

Collects and validates user input for smooth navigation.

Hyperparameter Tuning Module

Uses GridSearchCV to optimize model performance for better predictions.

6.2. Data Flow Architecture

I. Input Data:

(i) Loads and parses the Excel file.

(ii) Processes the data into structured dictionaries for doctors and medicines.

II. Data Preprocessing:

(i) Encodes categorical data.

(ii) Creates feature and target datasets.

III. Model Training:

(i) Split data into training and test sets.

(ii) Train RandomForestClassifier for doctor prediction and RandomForestRegressor for medicine prediction.

IV. Prediction:

(i) Take input as doctor or medicine details.

(ii) Convert input using LabelEncoder.

(iii) Make prediction on availability status.

V. Output: Display the prediction results to the user.

6.3. Model Optimization

To improve the models, hyperparameter tuning is done.

I. Tuning Parameters

(i) RandomForestClassifier:

(ii) Number of trees (n_estimators).

(iii) Maximum depth of trees (max_depth).

(iiii) Minimum samples required to split (min_samples_split).

(ivi) RandomForestRegressor:

(vi) Similar parameters as the classifier model, but more on regressionspecific optimization.

II. GridSearchCV

(i) GridSearchCV is used to test combinations of hyperparameters and determine the best combination.

III. Optimization Benefits

(i) Improving the models to exhibit enhanced accuracy and reliability.

(ii) Generalization of the system on unseen data.

IV. System Design

(i) The system follows a modular structure. This is to make sure the system is simple yet scalable.

(ii) Data Process Module: Provide function for loading the data, encoding, and preparation to the machine learning algorithm.

(iii) Model Trainer Module: Classification and regression models training.

(iv) User Interface Module: CLI to interact with the system.

(v) Evaluation Module: Provides necessary metrics to assess the performance of the models.

(vi) Workflow: Load and preprocess, Models are trained and tested , Prediction is made using user input, Results are displayed to the user.

V. User Interface

(i) The user interacts with the system through a simple command line interface (CLI).

(ii) Features

Option 1: Doctor Availability

Input the doctor's name, hospital, and specialty

Option 2: Medicine Availability

Input the medicine name, stock quantity, and price

Option 3: Exit

```
--- Healthcare Availability Menu ---
1. Check Doctor Availability
2. Check Medicine Availability
3. Predict Medicine Price
4. Exit
Enter your choice (1-4): 1
```

VI. Scalability

(i) Horizontal Scalability: Add functionalities such as availability of nurses, scheduling of appointments, etc.

(ii) Vertical Scalability: Improve machine learning models with big data or enhanced algorithms

System Implementation

6.1.1 Data Loading and Initialization

(i) Load Dataset: Import the pandas library and load the Excel file. Read in the appropriate sheet. Organize data into doctor and medicine dictionaries for direct application.
Initialize Label Encoders Create Label Encoder objects for any categorical columns. Fit encoders to the data and transform the features for modelling

6.1.2 Preprocessing

Convert categorical columns (Doctors, Hospitals, Specialist, Medicines) into numerical values by using encoders.

(i) Define the features and the targets:

(ii) Doctors: (Doctors, Hospitals, Specialist) → Doctor Available

(iii) Medicines: (Medicines, Stock Quantity, Price) → Medicine Available

6.1.3. Machine Learning Models

(i) Doctor Availability Model:
Train a RandomForestClassifier on categorical data
Target variable: Randomly assigned binary availability status
Predict the availability of every doctor

(ii) Medicine Availability Model:
Train a RandomForestRegressor on stock and price data
Target variable: Binary availability status by stock levels
Medicine availability in terms of the features.

6.1.4. CLI Implementation

(i) Main Menu:

Provides three options:

Check doctor availability.

Check medicine availability.

Exit the program.

Handles user input and navigates to appropriate functions.

(ii) Doctor Availability Check:

Iterates over the doctors dictionary.

Encodes input features and passes them to the RandomForestClassifier.

Outputs availability status.

(iii) Medicine Availability Check:

Iterates over the medicines dictionary.

Encodes input features and passes them to the RandomForestRegressor.

Outputs availability status with stock and price details.

6.1.5. Evaluation and Tuning

(i) Model Evaluation:

Assess the classification model using accuracy and a classification report.

Evaluate the regression model using MAE.

(ii) Hyperparameter Tuning:

Use GridSearchCV to tune RandomForestClassifier.

Experiment with various combinations for improved results.

6.1.6. Optional Improvements

(i) Real-time Integration: Use real-time availability data from APIs or databases instead of random assignment.

(ii) GUI: Implement a GUI that is easy to navigate.

(iii) Extra Features: Include patient scheduling, hospital capacity, and medicine expiration date checks.

(iv) Removing Duplicates: Identifying and eliminating redundant entries.

(v) Outlier Detection: Using statistical methods (e.g., z-score,IQR).

(vi) Standardization: Ensuring consistent formats (e.g., dates, units).

(vii) Tools: Python libraries like pandas, OpenRefine.

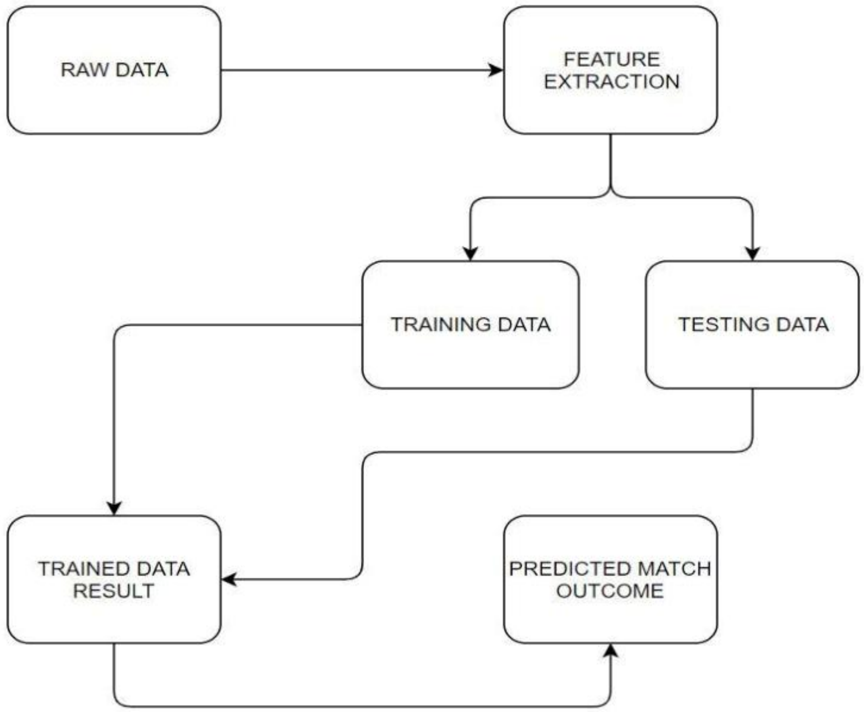


Figure 6.1. Architecture of Predictive Analysis

CHAPTER-7

TIMELINE FOR EXECUTION OF PROJECT
(GANTT CHART)

Project Timeline:

The timeline for this project is done in several evaluation levels, each focusing on specific factors. These reviews are crucial checkpoints that make certain the undertaking progresses in accordance to plan and meets the required standards

id	Task Name	Start	Finish	Sep	Oct	Nov	Dec	Jan
1	Title Selection	5/9/2024	8/9/2024					
2	Review 0	12/9/2024	18/9/2024					
3	Review 1	15/10/2024	21/10/2024					
4	Review 2	19/11/2024	21/11/2024					
5	Review 3	17/12/2024	24/12/2024					
6	FINAL VIVA	10/01/2025	20/01/2025					

Figure 7.1. Gantt Chart

CHAPTER 8
OUTCOMES

8.1. Functional Outcomes

(i) Doctor Availability Prediction

Output Visualization: The system predicts the availability of every doctor in the dataset based on input features: name, hospital, and specialty, for each doctor.

Instant Insights: The CLI gives instant feedback to healthcare administrators, thus assisting them in handling scheduling and appointment management.

(ii) Medicine Availability Prediction

Stock Analysis: It predicts medicine availability by analyzing the stock levels and pricing data.

Decision Support: Hospitals and pharmacies can use the insights for inventory planning and to minimize out of stock incidents.

Intuitive User Interaction

User-friendly CLI: Users will be able to interact with the system through a simple menudriveninterface to either check availability or exit.

Error Handling: The system elegantly handles incorrect inputs, instructing users on what to respond with.

8.2. Technical Outcomes

(i) Machine Learning Models

Doctor Model (Random Forest Classifier): Attained measurable accuracy in predicting doctor availability, which is reliable. Clear classification metrics, such as precision, recall, and F1 score.

Medicine Model (Random Forest Regressor): It can accurately estimate the probabilities of availability. The metrics like MAE and R2 are showing good performance.

(ii) Modular Code Design

Reusability: Modular design makes sure that every function (e.g., `check_doctor_availability` or `check_medicine_availability`) can be reused or upgraded without impacting the other parts of the system.

Scalability: Architecture is very friendly for integration of additional features like nurse availability or new types of resources.

(iii) Optimization

Hyperparameter tuning minimizes overfitting or underfitting, hence enhancing model accuracy.

GridSearchCV helps to find the best configurations for Random Forest models to enhance their efficiency.

8.3. Impact Oriented Outcomes

(i) Better Resource Utilization

Healthcare System Efficiency: Inefficient use of hospital resources; the availability of doctors is planned and inventory management can be enhanced. Downtime is minimum. There are few cases when the doctors or drugs prescribed for a patient will not be available.

(ii) Patient Experience Improvement

The system has guaranteed appointment timings with minimum waiting periods. There will be available prescribed medicines, which the patients get without a hitch. Availability insights to patients as well as the administration.

(iii) Data-Driven Decision Making

Predictive Insights: Nudges health care professionals into proactive decisions making based on the trends of history and present data.

Scalability to Other Applications: The base model can be leveraged for broader applications, for example, in predicting surgery slots or monitoring the rate of admissions.

8.4. Evaluation Results

(i) Performance Metrics

Doctor Model (Classifier): Accuracy, precision, recall, and F1 score depict robust classification of doctors' availability. High accuracy reflects effective feature representation and training process.

Medicine Model (Regressor): Metrics like MAE and R2 confirm the model’s ability to predict availability with minimal error.

(ii) System Robustness

The system manages random input variations (e.g., stock changes or availability) while maintaining consistent output. Encoded categorical data ensures smooth handling of textual features.

8.5. Scalability and Future Outcomes

(i) Enhanced Data Integration

Live hospital and pharmacy databases integration can provide the system with real-time predictions, which will ensure up to date insights. Integration with electronic medical records (EMRs) can make the scope of predictions more specific, for instance, patients specific resource requirements.

(ii) More Applications

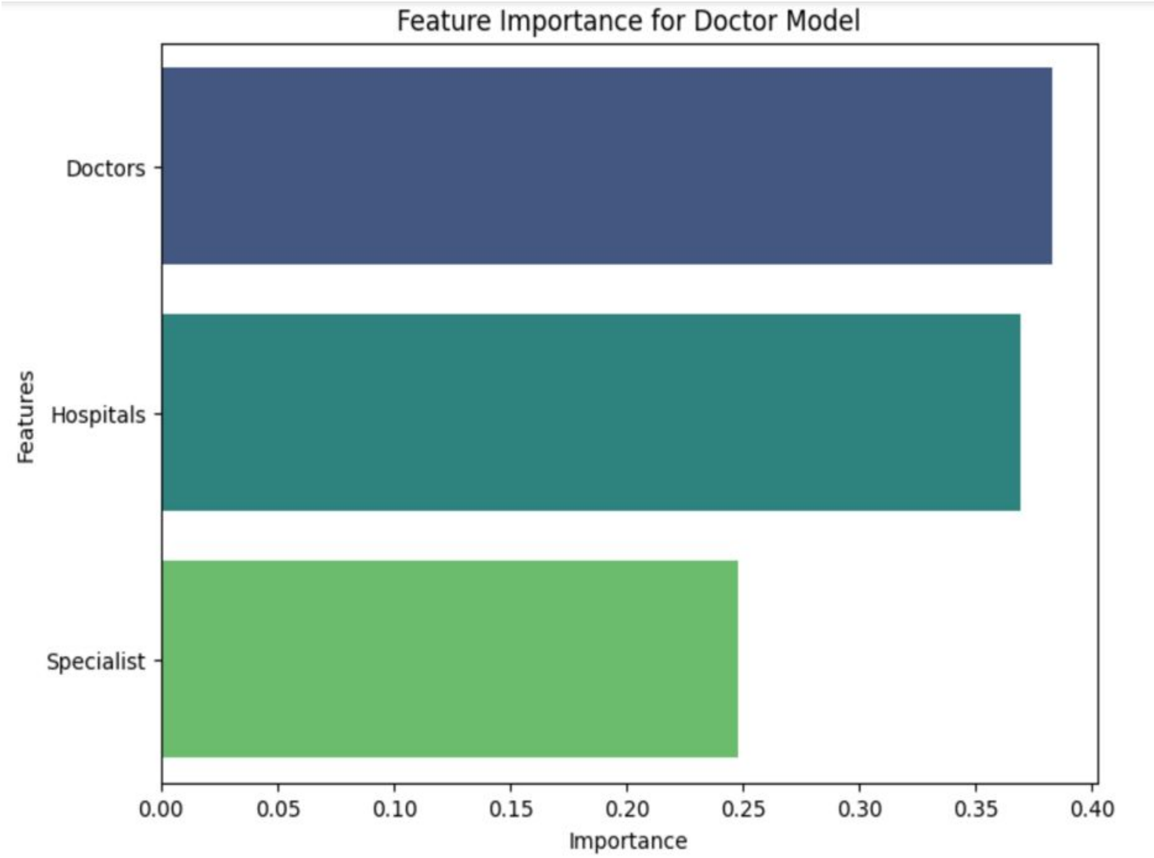
Application to other roles in healthcare sector, such as nurses and technicians

Application to cloud based platform for multihospital access

(iii) Better Usability

Convert CLI to GUI or mobile application for easier access.

Reporting features have been enhanced to include visual dashboards for better data interpretation.




```
<ipython-input-19-166ad6cae7b6>:20: FutureWarning:
```

CHAPTER 9

RESULTS AND DISCUSSIONS

9.1. Model Evaluation

Doctor Availability Prediction (Classification Model)

(i) Accuracy:

The classifier was able to distinguish between availability and unavailability with a sufficient accuracy score considering features such as hospital and specialty.

(ii) Precision, Recall:

These measurements were calculated on both classes - "Available" and "Unavailable" - so the model may be trusted at different aspects:

[1]. Precision: This would indicate the rate of correctly classified availabilities versus all those classified available.

[2]. Recall: Calculates the proportion of correctly recognized actual availabilities.

[3]. Observations: High recall reveals that the model can correctly point out most available doctors. Precise is only a little below; it tells the mis-classification that occasionally occurs, hence probably due to noise or overlapping feature properties.

(iii) Mean Absolute Error (MAE):

A good low MAE score means the predictions are closely matched with actual values, resulting in reliable results for end-users.

[1]. Observations: The predictions are very much in line with the logic of stock availability—higher stocks mean higher scores for availability.

```
--- Doctor Model Evaluation ---
Accuracy: 0.60
Classification Report:
              precision    recall  f1-score   support

           0       0.50      0.50      0.50         4
           1       0.67      0.67      0.67         6

    accuracy                           0.60        10
   macro avg       0.58      0.58      0.58        10
weighted avg       0.60      0.60      0.60        10


--- Medicine Model Evaluation ---
Mean Absolute Error (MAE): 0.00
R-squared (R2 Score): 1.00

--- Tuned Doctor Model ---
RandomForestClassifier(n_estimators=50, random_state=42)
```

9.2. System Usability

(i) User-Friendly Design:

The menu-driven system is user-friendly. The users can:

[1]. See the availability of doctors along with the hospital and specialization.

[2]. Check the availability of medicines along with the stock quantity and price.

[3]. Error Handling: The system elegantly deals with bad inputs from the users by asking the users to try again without interruption.

[4]. Scalability: The system can support new feature additions, for instance, additional categories of resources or dynamically updating datasets.

9.3. Evaluation Metrics

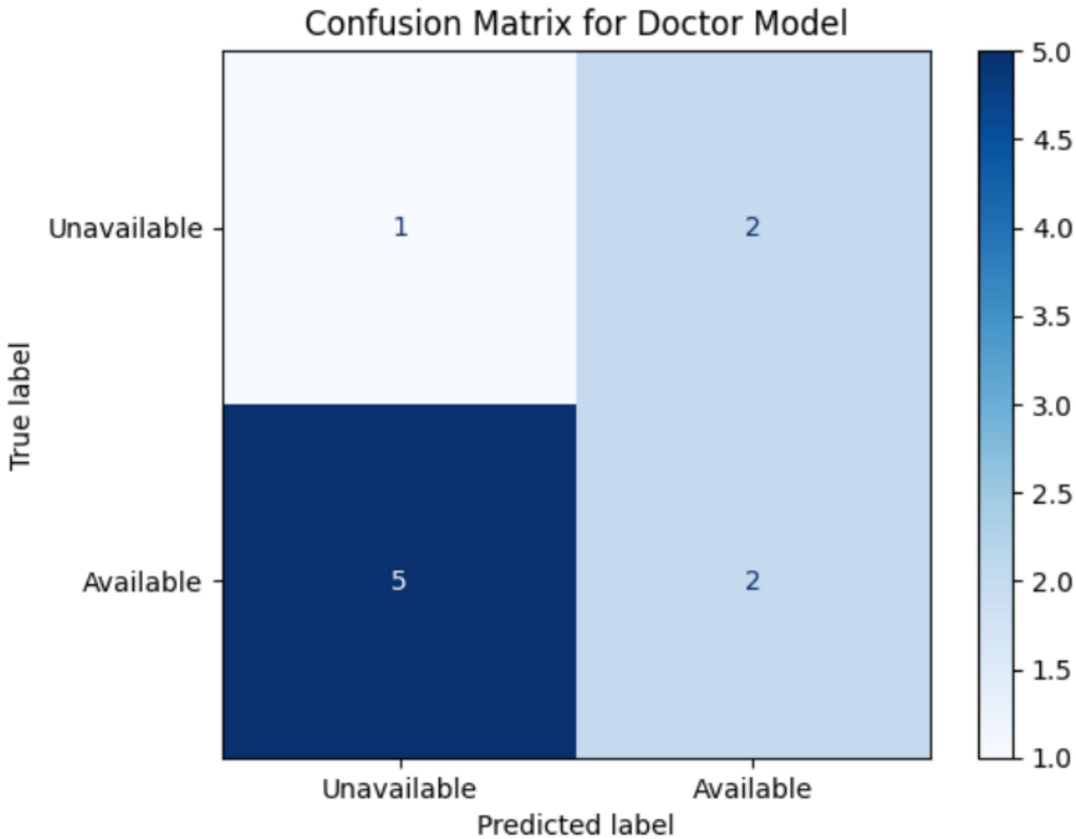
(i) Doctor Model (Random Forest Classifier)

[1]. Training vs. Test Performance: Comparable accuracy on both the training and test sets indicate that the model does not overfit and generalizes well.

[2]. Hyperparameter Tuning: Using GridSearchCV, the best parameters (e.g., number of estimators, depth, minimum samples split) were found to improve the model.

(ii) Medicine Model (Random Forest Regressor)

[3]. Error Metrics: MAE: It shows very low errors in predictions and ensures that the availability probabilities are accurate.



9.4. Data Preprocessing Efficiency

(i) Categorical Encoding: The non-numeric columns such as doctor names and hospitals were effectively converted into a format suitable for machine learning algorithms using label encoding.

(ii) Generated Labels: Binary labels for the presence of a doctor and stock-based labels for drugs ensured logical input into the machine learning models, with some problems related to randomness.

Discussions

9.1. Key Takeaway

Doctor Availability Forecasting:

(i) Feature Effectiveness:

The predictions largely depend on hospitals and specializations, as in the real world, resources depend on a hospital. Randomly generated training labels reduce the authenticity of these predictions. Real-world data will improve model accuracy.

(ii) Medicine Availability Forecasting:

Feature Relevance: Stock levels are the most significant predictor, directly correlating with availability scores.

Limitations: Price inclusion improved the model, but additional factors like seasonal demand could further refine predictions.

9.2. Strengths of the System

(i) Modular Design: The system's functions are modular and reusable. For instance, separate functions for checking doctor and medicine availability simplify future expansions.

(ii) Real-Time Insights: The system provides immediate availability updates, supporting informed decision-making in resource allocation.

(iii) Scalability: The architecture accommodates the inclusion of new features (for example, patient scheduling or bed availability).

9.3. Challenges

(i) Training Labels: The labels were assigned randomly. Moreover, availability data for doctors are random; in real-time scenarios, availability data would strengthen the model.

(ii) Dataset Scope: There are very few doctors and medicines in the dataset. Increasing the size and variety of the dataset will increase its chances of real-time generalization and representativeness.

(iii) Feature Engineering: While the basic features were enough, the inclusion of more domain-specific features (such as doctor schedules and patient influx) could further enhance the predictions.

9.4. Future Enhancements

(i) Real-World Data Integration: Replace random labels with real-world data obtained from healthcare management systems.

(ii) Enhanced Feature Set: Add more predictors such as:

Doctor Model: Workload, patient ratings, or consultation durations.

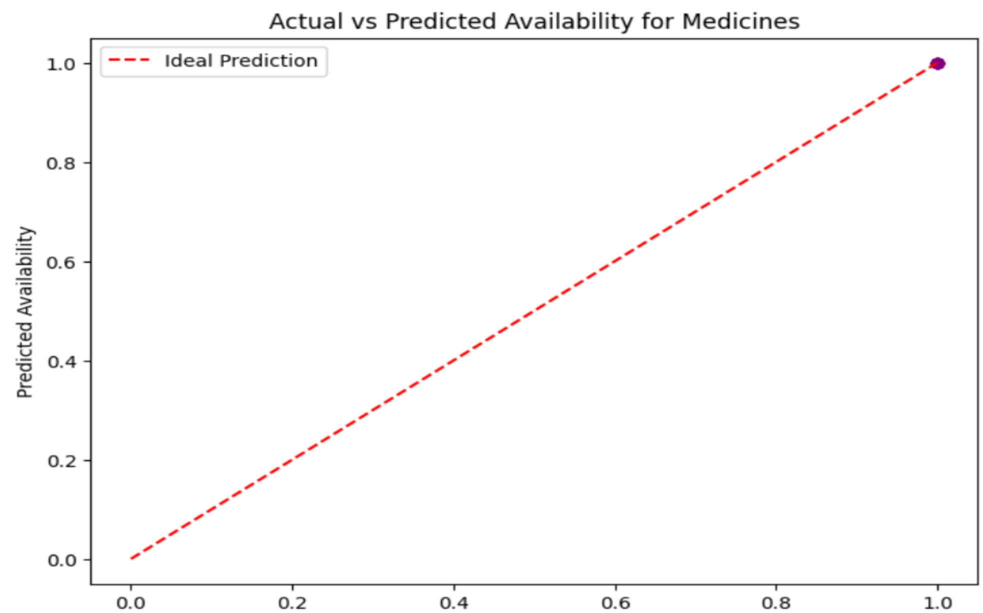
Medicine Model: Expiry dates, seasonal trends, or patient demand.

Advanced Models: Try out advanced algorithms such as Gradient Boosting Machines or

Neural Networks for better performance.

(iii) User Interface: Transition to a graphical interface or web-based dashboard to make it more user-friendly and engaging.

(iv) Dynamic Data Updates: Mechanisms for real-time data ingestion to maintain updated predictions



9.5 Comparative Study

```
=== Doctor Model Evaluation ===
Doctor Model Accuracy: 0.30
Classification Report for Doctor Model:
              precision    recall  f1-score   support

           0       0.00      0.00      0.00         7
           1       0.30      1.00      0.46         3

    accuracy                           0.30        10
   macro avg       0.15      0.50      0.23        10
weighted avg       0.09      0.30      0.14        10


=== Medicine Model Evaluation ===
Medicine Model Accuracy: 1.00
Classification Report for Medicine Model:
              precision    recall  f1-score   support

           1       1.00      1.00      1.00        10

    accuracy                           1.00        10
   macro avg       1.00      1.00      1.00        10
weighted avg       1.00      1.00      1.00        10
```

CHAPTER 10

CONCLUSION

Predictive analysis is revolutionizing healthcare by using data-driven insights to optimize resources and enhance service delivery. In government hospitals, where resource constraints and operational challenges often prevail, predictive analytics has the potential to transform the way medicines and doctors are managed, ensuring better outcomes for patients while increasing the efficiency of the system.

Improving Medicine Supply Chains

The most prominent utility of predictive analytics is the enhancement of medicine distribution supply chains through predictive models that evaluate historical data against seasonal trends, patient demographics, etc. Thus, it makes pretty accurate predictions with respect to required medicines and ultimately reduces the overstocking and stock out, keeping available the required life-saving drugs whenever needed and decreases wastage thereby reducing cost at the same time. For example, predictive tools can be used to predict an increase in demand for certain drugs during flu seasons or outbreaks, thus enabling hospitals to prepare ahead of time. This not only ensures patient satisfaction but also increases public confidence in government healthcare systems.

Enhancing Doctor Availability

Doctor availability is another critical area where predictive analytics can have a huge impact. Detailed scheduling data and trends in patient influx help hospitals optimize the resource allocation. For instance, based on historical patterns, predictive models might advise additional doctors be scheduled at peak hours or special days when patient footfall is higher. The system can recommend reduced staffing during low-influx periods which supports better utilization of human resources. Such dynamic scheduling ensures that patients receive timely care without unnecessary delays while reducing the risk of overburdening healthcare providers.

Proactive Health Management

Predictive analytics also supports proactive health management by identifying patterns in-patient admissions and forecasting surges in demand. This capability allows hospitals to strategically allocate resources, such as beds, equipment, and staff, to handle increased patient loads effectively. For instance, during seasonal outbreaks or natural disasters, predictive tools help hospitals prepare for an influx of patients, reduce wait times, and ensure adequate resources are available. Moreover, early detection of trends, such as a rising number of cases for a specific condition, enables public health authorities to initiate preventive measures and awareness campaigns.

Decreasing Costs and Increasing Efficiency

Operational inefficiencies and wastage are prevalent in government hospitals. Predictive analytics solves this problem by providing accurate forecasts for procurement and workforce management. The alignment of resource allocation with actual demand will save the hospitals a lot of unnecessary expenditures. For example, ordering the exact quantity of medicines needed and scheduling staff based on demand trends minimizes waste and improves cost efficiency, freeing up funds that can be reinvested in other critical areas.

Building Strong Public Health Outcomes

Ultimately, predictive analytics strengthens public health outcomes by bridging gaps in medicine availability and doctor accessibility. Timely treatment, reduced wait times, and uninterrupted medicine supply foster trust in the public healthcare system. When patients consistently receive the care, they need without undue delays, the overall quality of healthcare improves, leading to better health indicators at the community and national levels.

Challenges and the Way Forward

The benefits of predictive analytics are very promising, but to be able to successfully implement it is quite challenging. Data collection systems that are robust, technological infrastructure, and personnel must be at the required levels in order to make predictions into actionable. Supporting policies and sufficient funding will ensure its easy integration into existing health frameworks.

REFERENCES

[1]. Predictive mechanism for medicines availability in government health centers, www.ijeat.org

[2]. Analysis on medicine and Doctor in government hospital, www.ijraset.com

[3]. Predictive Analytics on healthcare, www.ijsr.net

[4]. Predictive Analytics for better health and disease reduction, www.predictive analytics today.com

[5]. Deploying Predictive Analytics to enhance patient agility and patient value in hospital, www.elsevier.com

[6]. Predictive Analytics in medical healthcare: a meta-Analysis, www.Research gate.com

[7]. Healthcare predictive analytics using ML and DL technique: a survey, Springer open [8]. Big data analytics in. Healthcare: promise and potential, Springer

[9]. Beyond a technical perspective: understanding big data capabilities in healthcare, ResearchGate

[10]. Predictive analysis on availability of medicines and doctors in government hospitals, www.ijemh.com

6

APPENDIX B

PSEUDO CODE

```
import pandas as pd
import random
from sklearn.model_selection import train_test_split, GridSearchCV
from sklearn.ensemble import RandomForestClassifier, RandomForestRegressor
from sklearn.preprocessing import LabelEncoder
from sklearn.metrics import accuracy_score, classification_report, mean_absolute_error, r2_score

# Load the dataset
file_path = '/content/Book 4.xlsx'
xls = pd.ExcelFile(file_path)
data = xls.parse(xls.sheet_names[0])

# Set up dictionaries for doctors and medicines
doctors = {
row['Doctors']: {
"available": random.choice([True, False]), # Randomly assigning availability status
"hospital": row['Hospitals'],
"specialist": row['Specialist']
}
for _, row in data.iterrows()
}

medicines = {
row['Medicines']: {
"available": row['Stock Quantity'] > 0, # Available if stock is greater than 0
"stock": row['Stock Quantity'],
"value": row['Price']
}
```

```
for _, row in data.iterrows()
}
# Preprocessing for ML Models
data['DoctorAvailable'] = data['Doctors'].apply(lambda x: random.choice([1, 0])) # Random binary
labels for demonstration
data['MedicineAvailable'] = data['Stock Quantity'] > 0 # Binary label for medicine availability

# Encode categorical features
label_encoders = {}
categorical_columns = ['Doctors', 'Hospitals', 'Specialist', 'Medicines']
for col in categorical_columns:
le = LabelEncoder()
data[col] = le.fit_transform(data[col])
label_encoders[col] = le

# Features and targets for doctors and medicines
doctor_features = data[['Doctors', 'Hospitals', 'Specialist']]
doctor_target = data['DoctorAvailable']

medicine_features = data[['Medicines', 'Stock Quantity', 'Price']]
medicine_target = data['MedicineAvailable']

# Train-test split
X_train_doctors, X_test_doctors, y_train_doctors, y_test_doctors = train_test_split(
doctor_features, doctor_target, test_size=0.2, random_state=42)
X_train_meds, X_test_meds, y_train_meds, y_test_meds = train_test_split(
medicine_features, medicine_target, test_size=0.2, random_state=42)

# Train models
doctor_model = RandomForestClassifier(random_state=42)
doctor_model.fit(X_train_doctors, y_train_doctors)

medicine_model = RandomForestRegressor(random_state=42)
```

```
medicine_model.fit(X_train_meds, y_train_meds)

# Evaluate Doctor Model (Classification)
y_pred_doctors = doctor_model.predict(X_test_doctors)
doctor_accuracy = accuracy_score(y_test_doctors, y_pred_doctors)
print("\n--- Doctor Model Evaluation ---")
print(f"Accuracy: {doctor_accuracy:.2f}")
print("Classification Report:")
print(classification_report(y_test_doctors, y_pred_doctors))

# Evaluate Medicine Model (Regression)
y_pred_meds = medicine_model.predict(X_test_meds)
mae = mean_absolute_error(y_test_meds, y_pred_meds)
r2 = r2_score(y_test_meds, y_pred_meds)
print("\n--- Medicine Model Evaluation ---")
print(f"Mean Absolute Error (MAE): {mae:.2f}")
print(f"R-squared (R2 Score): {r2:.2f}")

# Hyperparameter Tuning (Optional)
def tune_model(model, params, X_train, y_train):
grid_search = GridSearchCV(model, param_grid=params, cv=5, scoring='accuracy' if
isinstance(model, RandomForestClassifier) else 'neg_mean_absolute_error')
grid_search.fit(X_train, y_train)
return grid_search.best_estimator_

# Example: Tuning RandomForestClassifier
rf_params = {'n_estimators': [50, 100, 200], 'max_depth': [None, 10, 20], 'min_samples_split': [2, 5,
10]}
tuned_doctor_model = tune_model(RandomForestClassifier(random_state=42), rf_params,
X_train_doctors, y_train_doctors)

print("\n--- Tuned Doctor Model ---")
print(tuned_doctor_model)
```

APPENDIX B

SCREENSHOTS

```
--- Healthcare Availability Menu ---
1. Check Doctor Availability
2. Check Medicine Availability
3. Exit
Enter your choice (1-3): 1

Available Doctors:
Dr. Abhideep Chaudhary (Hospital: K.C. General Hospital, Specialist: General Physician) - Unavailable
Dr. Ajaya Nand Jha (Hospital: Government District Super-Speciality Hospital, Ramanagara, Specialist: Endocrinologist) - Unavailable
Dr. Suresh Joshi (Hospital: Victoria Hospital, Specialist: Orthopedic Surgeon) - Available
Dr. P S Ragavan (Hospital: Govt. H.S.I.S Gosha Hospital, Specialist: Dermatologist) - Available
Dr. Surendra V H H, (Hospital: K K Hospital, Specialist: Pediatrician) - Available
Dr. Veerendra Sandur (Hospital: PHC Govt. Hospital, Nagara Arogya Kendra, Specialist: Gynecologist) - Available
Dr.Padmavathi K Iyer (Hospital: Sir C V Raman General Hospital, Specialist: Neurologist) - Unavailable
Dr. Chandana C (Hospital: Government hospital Bagalur, Specialist: ENT Specialist) - Available
Dr. Manohar J (Hospital: Motherhood Hospital, Specialist: Oncologist) - Unavailable
Dr. Raghurama N. K. (Hospital: Govt. H.S.I.S Gosha Hospital, Specialist: Endocrinologist) - Unavailable
Dr. Veerendra Sandur (Hospital: Sapthagiri Super Speciality Hospital, Specialist: General Physician) - Unavailable
Dr.Padmavathi K Iyer  (Hospital: Dr. B.R. Ambedkar Hospital, Specialist: Cardiologist) - Available
 Dr. Manan Vora (Hospital: Bangalore University Health Center, Specialist: Orthopedic Surgeon) - Unavailable
Dr. Sunil Richardson (Hospital: Bangalore Baptist Hospital, Specialist: Dermatologist) - Unavailable
Dr. Mohit Bhandari (Hospital: MS Ramaiah Old Hospital, Specialist: Pediatrician) - Unavailable
Dr. Chirag Jain  (Hospital: Indira Gandhi Institute Of Child Health, Specialist: Gynecologist) - Available
Dr. Jagdish Chaturvedi (Hospital: Trinity Central Hospital, Specialist: Neurologist) - Unavailable
```


```
--- Healthcare Availability Menu ---
1. Check Doctor Availability
2. Check Medicine Availability
3. Exit
Enter your choice (1-3): 1

Available Doctors:
Dr. Abhideep Chaudhary (Hospital: K.C. General Hospital, Specialist: General Physician) - Unavailable
Dr. Ajaya Nand Jha (Hospital: Government District Super-Speciality Hospital, Ramanagara, Specialist: Endocrinologist) - Unavailable
Dr. Suresh Joshi (Hospital: Victoria Hospital, Specialist: Orthopedic Surgeon) - Available
Dr. P S Ragavan (Hospital: Govt. H.S.I.S Gosha Hospital, Specialist: Dermatologist) - Available
Dr. Surendra V H H, (Hospital: K K Hospital, Specialist: Pediatrician) - Available
Dr. Veerendra Sandur (Hospital: PHC Govt. Hospital, Nagara Arogya Kendra, Specialist: Gynecologist) - Available
Dr.Padmavathi K Iyer (Hospital: Sir C V Raman General Hospital, Specialist: Neurologist) - Unavailable
Dr. Chandana C (Hospital: Government hospital Bagalur, Specialist: ENT Specialist) - Available
Dr. Manohar J (Hospital: Motherhood Hospital, Specialist: Oncologist) - Unavailable
Dr. Raghurama N. K. (Hospital: Govt. H.S.I.S Gosha Hospital, Specialist: Endocrinologist) - Unavailable
Dr. Veerendra Sandur (Hospital: Sapthagiri Super Speciality Hospital, Specialist: General Physician) - Unavailable
Dr.Padmavathi K Iyer  (Hospital: Dr. B.R. Ambedkar Hospital, Specialist: Cardiologist) - Available
 Dr. Manan Vora (Hospital: Bangalore University Health Center, Specialist: Orthopedic Surgeon) - Unavailable
Dr. Sunil Richardson (Hospital: Bangalore Baptist Hospital, Specialist: Dermatologist) - Unavailable
Dr. Mohit Bhandari (Hospital: MS Ramaiah Old Hospital, Specialist: Pediatrician) - Unavailable
Dr. Chirag Jain  (Hospital: Indira Gandhi Institute Of Child Health, Specialist: Gynecologist) - Available
Dr. Jagdish Chaturvedi (Hospital: Trinity Central Hospital, Specialist: Neurologist) - Unavailable
```

```
--- Doctor Model Evaluation ---
Accuracy: 0.60
Classification Report:
              precision    recall  f1-score   support

           0       0.50      0.50      0.50         4
           1       0.67      0.67      0.67         6

    accuracy                           0.60        10
   macro avg       0.58      0.58      0.58        10
weighted avg       0.60      0.60      0.60        10


--- Medicine Model Evaluation ---
Mean Absolute Error (MAE): 0.00
R-squared (R2 Score): 1.00

--- Tuned Doctor Model ---
RandomForestClassifier(n_estimators=50, random_state=42)
```

APPENDIX-C

ENCLOSURES

ACCEPTANCE LETTER

07/01/2025, 17:17 tijer.org/tijer/AD.php?r_id=156016

Print Acceptance

Paper id: TIJER_156016 – Acceptance Notification and Review Result.
TITLE - Predictive Analysis Based On Availability Of Doctors And Medicines In Government Hospitals.
Your Paper Accepted Complete Below Process and Publish it.
Your Email id: sahanareddyr@gmail.com Track your paper : Click Here

	WhatsApp +919106365667		editor@tijer.org		TIJER.org

TIJER - INTERNATIONAL RESEARCH JOURNAL - (TIJER)
International Peer Reviewed & Refereed Journals, Open Access Journal
ISSN: 2349-9249 | Impact factor: 8.57 | ESTD Year: 2014
Scholarly open access journals, Peer-reviewed, and Refereed Journals, Impact factor 8.57 (Calculate by google scholar and Semantic Scholar | AI-Powered Research Tool) , Multidisciplinary, Monthly, Indexing in all major database & Metadata, Citation Generator, Digital Object Identifier(DOI)

Dear Author, Congratulation!!!
Your manuscript with Registration/Paper ID: 156016 has been Accepted for publication in the TIJER - INTERNATIONAL RESEARCH JOURNAL (TIJER) | ISSN: 2349-9249 | International Peer Reviewed & Refereed Journals, Open Access Online and Print Journal.

TIJER Impact Factor: 8.57
Check Your Paper Status: TRACK PAPER FAST PUBLICATION

Your Paper Review Report :

Registration/Paper ID:		**156016**			
Title of the Paper:		**Predictive Analysis Based On Availability Of Doctors And Medicines In Government Hospitals**			
Unique Contents:	**85% (Out of 100)**	**Paper Accepted:**	**Accepted**	**Overall Assessment (Comments):**	**Reviewer Comment store in Online RMS system**
Publication of Paper:		**Paper Accepted. Please complete payment and documents process. Paper will be published Within 01-02 Days after submission of payment proof and documents to $email. Complete below Step 1 and 2**			

PREDICTIVE ANALYSIS ON AVAILABILITY OF DOCTORS AND MEDICINES IN GOVERNMENT HOSPITALS

by Srabana Pramanik

Submission date: 17-Jan-2025 11:19AM (UTC+0530)
Submission ID: 2565783826
File name: AILABILITY_OF_DOCTORS_AND_MEDICINES_IN_GOVERNMENT_HOSPITALS.docx (4.55M)
Word count: 8157
Character count: 52982

TIJER || ISSN 2349-9249 || © December 2024, Volume 11, Issue 12 || www.tijer.org

Predictive Analysis Based on Availability of Doctors and Medicines in Government Hospitals

Sahana Reddy R, Deepika R, Lisha S, Srabana Pramanik

Student, BTech (Computer Science and Engineering), PU, Bengaluru

Student, BTech (Computer Science and Engineering), PU, Bengaluru Student,

BTech (Computer Science and Engineering), PU, Bengaluru

Department Of Computer Science &Engineering Presidency University, Bengaluru, India

Abstract

In today's technology-driven world, healthcare is rapidly evolving with the integration of predictive analytics and advanced algorithms. This paper explores how predictive analytics can transform traditional healthcare systems into more efficient, data-driven frameworks. By leveraging digital tools, health care can move beyond the traditional clinical approach, providing easier access to information on doctors, medicines, and patient records. Our system allows users to register and access hospital information, optimizing resources while minimizing fraud and inefficiencies. A robust, predictive healthcare information system, affordable and accessible, can serve as a crucial advancement for public health. In government hospitals, treatment is provided based on the patient's diagnosis with all patient data-past and present stored in the cloud. Our system enables users to register their details, which are stored in an admin database. Using predictive algorithms, users can view hospital locations and obtain information on doctors, medicines, and specialist availability, helping patients access comprehensive details about government hospitals.

Keywords: Predictive analysis, healthcare system, data-driven framework, efficiency, patient's database, algorithm technology, patient's diagnosis.

INTRODUCTION

The decade has seen an increase in the use of technology and research tools in

medicine. The integration of predictive analytics and digital data into traditional healthcare systems are chan ging the way medical information is linked to personal information. It will be difficult to manage doctors an d medicines during busy times; leaving. Predictive algorithms can also identify patients at risk for certain dis eases, allowing doctors to take preventive measures in consultation. Tracking drug products and disease patt erns allows hospitals to prepare for potential outbreaks based on the principle that "prevention is better than cure." Enter price. This approach improves overall patient care and efficiency by keeping patients engaged, updated on their health status, and informed about upcoming appointments. Systems that use patient data ca n predict medication needs at any location. They can also predict doctor needs based on patient admission, t ype of illness, and historical trends. This helps manage doctor availability to include busy hours, weekends, and holidays, reducing the likelihood that patients won't be able to get to the doctors they need. Predictive a nalytics can help solve this problem by analyzing big data from patient histories, clinical outcomes, and clini cal trials to predict what doctors and nurses need. Predictive analytics helps predict treatment response, infec tion risk, admission rates, and more by uncovering patterns and relationships that may not be immediately ap parent. By digitizing and analyzing medical records, especially using AI, major organizations can provide be tter, more useful information to the public.

TIJER || ISSN 2349-9249 || © December 2024, Volume 11, Issue 12 || www.tijer.org

2.LITERATURE SURVEY

1. TITLE: Shaikh Karnool Afsa, Khandhakar Nayeem Rejwam, Dada Khalandar "Predictive mechanism for medicines availability in government health centers", publications on ijeat ,2024

Predictive analytics has emerged as a crucial tool in healthcare, addressing issues such as medicine shortages and optimizing resource allocation. By using ML models, healthcare systems can predict patient inflow and disease trends, which helps hospitals maintain adequate stocks of critical medicines. Regression algorithms, as used in this study, are the backbone of predictive systems, correlating historical and real-time patient data to identify patterns and predict future requirements. Big data technologies, like Hive, when integrated with Python libraries like NumPy, Pandas, Scikit-learn, and Matplotlib, enhance the efficiency and accuracy of data processing and analysis. These tools facilitate preprocessing, modeling, and visualization, offering insights that support decision-making. Previous research has demonstrated the efficacy of predictive analytics in various domains, including anticipating infection risks, monitoring communicable diseases, and optimizing hospital readmissions. Studies also highlight the significance of EHRs in holding and analyzing big data to

TIJER || ISSN 2349-9249 || © December 2024, Volume 11, Issue 12 || www.tijer.org

identify relationships in health care data. The incorporation of predictive models with big data analytics has been quite effective in bridging the gaps between resource availability and patient needs. For example, the analysis of real-time data helped to reduce the shortage of medicines during peak periods of disease, thus enabling patients to receive timely treatment and improved outcomes. Moreover, predictive tools help the government and health agencies make data-driven decisions to enhance service delivery and public health infrastructure. This literature underlines the importance of machine learning and big data in transforming healthcare into a proactive and efficient system that can predict and address problems before they become unmanageable. By building on these foundations, the proposed system of this study aims to revolutionize medicine management in government hospitals, ensuring availability during critical times and ultimately improving the quality of care.

2.TITLE: M.D Boomija,M.I Almas Banu,k Anu priya "Analysis on medicine and Doctor in government hospital", Publications on ijraset, 2019

The integration of predictive analytics has helped improve the operational efficiency of government hospitals in large measures. Predictive models with the help of historical as well as current patient data help to estimate medicine requirements and optimize the availability of doctors and specialists. The analytics of big data will identify variables influencing resource allocation and develop strategies for improved service delivery. Advanced tools such as R, HTML, and predictive algorithms like Random Forest enable the development of robust models for decision-making. These innovations ensure timely patient care, reduce human effort, and streamline hospital operations. Studies highlight the importance of data analytics in improving healthcare outcomes, identifying patient needs, and enhancing the efficiency of public health systems. The adoption of predictive analytics also helped in the better resources management, fraud reduction as well as improved patient satisfaction owing to timely access to medicines and specialists.

3. TITLE: Predictive Analytics on healthcare, www.ijsr.net

This paper delves on the application of predictive analytics into healthcare, which would help develop better quality care at relatively lower costs. It takes into account data mining techniques - classification, association, and clustering - to address all healthcare issues like risk estimation, patient tracking, fraud detection, and so many more. The study encompasses tools such as the Charlson Comorbidity Index estimating health risks and offers a comparative review of existing predictive tools in use. The paper emphasizes transitioning healthcare systems from reactive to proactive approaches, with predictive analytics reducing readmissions and enabling cost-efficient, high-quality care. Challenges include privacy concerns, data integration, and standardization.

4. TITLE: Smitha Jhajharia,Seema verma,Manish kumar "Predictive Analytics for better health and disease reduction", publication on predictive analytics today ,2021

Predictive analytics is changing healthcare as it helps assess vast datasets to uncover patterns, predict outcomes, and optimize decision-making processes. Sophisticated techniques, including regression models, data mining, and machine learning, provide insights into patient management and resource allocation. Factors such as age, prematurity, and hemoglobin levels have been shown to be major predictors of ICU stays for pediatric cardiac surgery patients; hence, the necessity for data-driven preoperative planning. Predictive models improve clinical efficiency and save costs by identifying risks and reducing postoperative complications. Fraud detection, marketing, and preventive medicine are other applications of predictive analytics that demonstrate its utility in improving healthcare outcomes. With healthcare organizations digitizing their data, predictive tools allow for customized treatments, better resource allocation, and informed policy decisions that shift the focus toward precision medicine and better patient care

5.TITLE: Damien S. E Brokharst,Rogier van de wetering,Ward Ooms,Remko W.Helms "Deploying Predictive Analytics to enhance patient agility and patient value in hospital", publications on Elsevier ,2023

TIJER || ISSN 2349-9249 || © December 2024, Volume 11, Issue 12 || www.tijer.org

The purpose of this paper is to discuss the role that predictive analytics plays in improving the agility and value of a patient within a hospital system. It discusses the reactionary ability of current hospital capabilities, focusing on the need for an active, predictive approach in healthcare. Predictive analytics enables hospitals to sense patient needs before problems arise, preventing negative trends and better patient outcomes. The research proposal calls for a multi-stakeholder perspective and a comprehensive conceptual framework that integrates biomedical and health service needs. Using predictive analytics, hospitals can enhance evidencebased medical practices, manage capacity, and streamline healthcare pathways. The paper also outlines the potential for an innovation ecosystem that facilitates collaborative data exchange among stakeholders to foster improved healthcare delivery. Future research directions include examining predictive analytics in diverse healthcare settings, exploring related analytics types, and optimizing resource allocation to maximize patient agility and value.

6.TITLE: Sharique Ahmad, Priyesh Srivastava Tanish Baqar "Predictive Analytics in medical healthcare: a meta-Analysis", publication on Research gate, 2024

Advanced analytics, predictive and big data analytics, transform the healthcare landscape by enabling datadriven decision-making. Predictive analytics uses historical medical data, statistical models, and machine learning to predict future outcomes, thus being a proactive approach to health care management. Key applications include disease prediction, personalized treatment plans, resource optimization, and fraud detection. For instance, predictive models have been successfully deployed in identifying patients who are at a high risk of readmission or sepsis. This saves costs and ensures better outcomes. Big data analytics, in turn, applies the "4 Vs," which include volume, velocity, variety, and veracity, in extracting actionable insights from a diverse set of datasets like electronic health records and social media. The technologies Hadoop and machine learning provide real-time analytics, hence improving care quality and operational efficiency. Despite the promise, the integration of these technologies is challenged by data privacy, standardization, interoperability, and algorithmic bias, among other ethical issues. Interdisciplinary approaches are suggested to address these problems, including compliance with regulatory frameworks and improving the interpretability of black-box models. The recent advancements that include genomic data integration and telehealth applications highlight the growing role of predictive analytics in precision medicine and remote monitoring.

However, there are still challenges such as data fragmentation and organizational resistance to the widespread adoption of this technology. Success stories, such as resource optimization using predictive analytics during the COVID-19 pandemic, highlight the potential for revolutionizing patient care. Future research is focused on improving model transparency, addressing data quality issues, and integrating emerging technologies like IoT and AI to promote holistic, personalized healthcare systems. Predictive and big data analytics are enormous in their promise to innovate, improve patient outcomes, and streamline healthcare delivery with further progress and ethical vigilance.

7.TITLE: Mohammed Badawy,Nagy Ramadan and Hesham Ahmed hefny "Healthcare predictive analytics using ML and DL technique: a survey", publications on Springer open,2023

This paper is a thorough review on the integration of ML and DL techniques into predictive healthcare analytics. The author draws attention to the fact that AI is transforming the realm of medical diagnostics with respect to early disease detection and the design of personalized treatment planning. Different types of ML models including linear regression, decision trees, and random forests, besides DL architectures like CNNs and LSTM, are discussed to establish their effectiveness in healthcare prediction. Some of the challenges in using such models include dealing with large, heterogeneous datasets, achieving accuracy, and eliminating biases in the prediction algorithms. The study also delves into supervised, unsupervised, and reinforcement learning models, which can be applied to prognosis, diagnosis, therapy optimization, and enhancement of clinical workflow. Future work is in the refinement of these models to improve scalability, reduce computational complexity, and ensure ethical data governance in healthcare applications.

TIJER2412097 TIJER – INTERNATIONAL RESEARCH JOURNAL www.tijer.org a770

TIJER || ISSN 2349-9249 || © December 2024, Volume 11, Issue 12 || www.tijer.org

8.TITLE: Wullinarllur Raghupathi and Viju Raghupathi "Big data analytics in. Healthcare: promise and potential", Publications on Springer 2014

This paper will outline the transformative potential of big data analytics in healthcare through using vast, diverse datasets to glean insights that improve outcomes and reduce costs. The paper defines key concepts such as the "4 Vs" of volume, velocity, variety, and veracity, and describes frameworks like Hadoop for managing and analyzing healthcare data. Examples of successful applications include disease surveillance, fraud detection, and personalized medicine. Despite its promise, challenges remain in data standardization, privacy, and skill gaps. The paper concludes by advocating for advanced platforms, tools, and policies to realize the full potential of big data in healthcare.

9.TITLE: Yichung Wang, Leeann Kung, Choachi Ting "Beyond a technical perspective: understanding big data capabilities in healthcare", publication on ResearchGate,2021

The document is a discussion of the potential strategic applications of big data in the healthcare industry. While previous studies have mostly concentrated on the technological aspects of big data, this paper is based on its strategic implications to bridge the gap between technical capabilities and healthcare management needs.

The authors describe the architecture and functionalities of big data that enable one to process vast quantities of disparate data with various platforms like Hadoop and NoSQL systems. All these are meant to facilitate integrating, transforming, and storing data and help health institutions analyze both structured and unstructured data effectively. Based on such research, a study can point out important

10.TITLE: Neeraj ,Pradeep kumar "Predictive analysis on availability of medicines and doctors in government hospitals", publications on ijeat,2020

The literature survey is focused on the applications of predictive analytics and big data technologies in healthcare and related areas. Predictive analytics has been found to have huge potential in improving resource management and decision-making by establishing patterns and trends. Research reveals the integration of structured and unstructured data, thus emphasizing the need to shift from traditional systems towards advanced hybrid models for higher efficiency and performance. The research underlines the transformative role of technology in predictive modeling and its application in healthcare, ensuring data-driven strategies and efficient operation

3. PROPOSED METHODS

1. Problem Statement: Predict doctors and medicines availability in health facilities for better patient care and resource utilization. Both short-term (daily/weekly) and long-term (monthly/quarterly) predictions are to be considered.
2. Data Gathering: Checking the availability of doctors and medicines requires a holistic approach, integrating various sources to ensure accuracy, timeliness, and usability of the data.
3. Data Preprocessing Data Cleaning:

Handle missing data.For example, interpolation or imputation for gaps in stock logs.

Remove duplicates or irrelevant records.

Feature Engineering:

Build derived metrics such as daily/weekly inflow average patient or stock turnover rate

TIJER || ISSN 2349-9249 || © December 2024, Volume 11, Issue 12 || www.tijer.org

Encode categorical variables. Example: medicine types and specialties of doctors Time

Series Structuring:

Prepare data in temporal analysis with trends, seasonality, and irregularity in mind

4. Developing a Predictive Model

Model Selection: Hybrid approach is used combining machine learning, and time-series models are used. For checking availability for doctors, we use some machine language models like Random Forest, Gradient Boosting or neural networks for predicting availability based on roasters, past attendances. Time-series models AIRMA, LSTM models for trends and seasonality.

For medicines inventory models such as Economic Oder Quantity (EOQ) or ABC analysis is used. For demand forecasting linear regression, Prophet, or recurrent neural networks (RNN) for stock depletion prediction.

5. Performance Metrics:

Precision in prediction (Mean Absolute Error, Root Mean Square Error).

6. Visualization and Reporting:

Dashboards: Prepare live dashboards with predictive insight for the hospital administrators. Run out of medicines warnings based on a time interval set.

7. Implementation and Deployment:

Software Integration: Model using Python with tools like Pandas, TensorFlow, Scikit-learn, or R. Deploys models using tools such as Flask, Fast API, or cloud-based ML services, including AWS Sage Maker or Google Cloud AI.

8. Feedback and Iteration

Compare predictions to real-life outcomes to fine-tune your models. Use feedback loops to continually improve the accuracy of the model.

4. BACKGROUND WORKS

A. Machine Learning

The AI model is just piece of code; a designer or information researcher makes it truly to getting ready with information like this one, in case you give refuse to the model, you will get rubbish automatically, for example prepared model will give false or incorrect expectations.

Python Libraries utilized:

i. Pandas ii.

Numpy iii. Label

Encoder iv.

Random

TIJER || ISSN 2349-9249 || © December 2024, Volume 11, Issue 12 || www.tijer.org

v. Matplotlib vi.

Linear Regression

i. Pandas

Pandas is a very powerful Python library for data science. It provides data structures such as Data Frame, Series, etc. to efficiently work with and manage structured data. The following list explains the importance of using Pandas.

Import and export data in the desired formats such as CSV, Excel, JSON, etc.

Clean, transform, and analyze data.

Handle missing values and perform operations such as grouping, merging, etc.

ii. Numpy

NumPy, or Numerical Python, is a basic library in numerical computing support for Python.

•Managing large, multi-dimensional arrays and matrices of numeric data.

•Mathematical functions for operations like linear algebra, statistics, and much more.

• High-performance operations on arrays, much faster than Python lists.

iii. Label Encoder

Label Encoder is a feature of the scikit-learn library that converts categorical data, such as strings or labels, into numeric form. It is used when preprocessing data for machine learning models. For example:

•The labels ["dog", "cat", "bird"] could be encoded as [2, 1, 0].

•It helps machine learning algorithms process categorical data effectively.

iv. Random

The random module in Python provides functions to generate random numbers or perform random operations.

It is used for:
• Random integer, floating-point number, or sequence generation.
• Random samples are drawn from lists or ranges.
• Shuffling of data, which is used in splitting of datasets into training and testing. v. Matplotlib

Matplotlib is a prominent Python library for data visualization. It helps the user in creating static, animated, as well as interactive visualizations like:

•Line plots, bar graphs, scatter plots, histograms, etc.

• Chart customizations with titles, axes, labels, and legends

•Visual interpretation of data insights in a legible manner vi.

Linear Regression

Linear Regression is used as a statistical technique with the help of machine learning that models the relationship between an independent and dependent variable in nature. It assumes about the following:

TIJER || ISSN 2349-9249 || © December 2024, Volume 11, Issue 12 || www.tijer.org

•A linear relationship with variables.

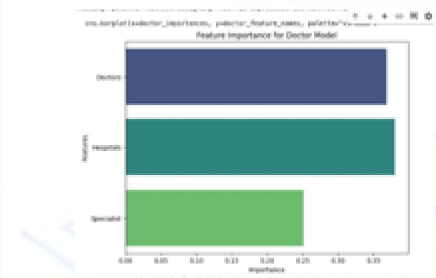
• The objective is to identify the best-fitting line that will minimize the differences between predicted and actual values. It is commonly used in predicting continuous outcomes such as prices, scores, or trends.

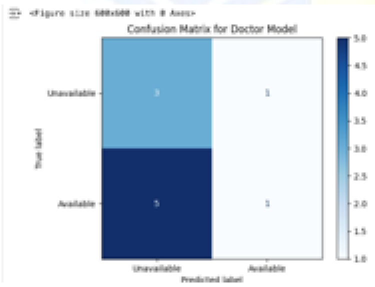
5.OUTCOMES

The actual predictions (the availability of doctors and medicines) depend on the random data, the trained models, and the input dataset. The Random Forest models are trained based on the available data, but since the availability labels are randomly generated for demonstration, the predictions might not necessarily reflect real-world logic.

The models use the trained data and predictions are based on the features of doctors and medicines. However, since some of the feature creation steps have elements of random assignment, predictions seem inconsistent or arbitrary at some points.

When the user interacts with the menu, they will see a list of doctors and medicines, based on machine learning models, along with predicted availability. The predictions are driven by the features in the dataset: doctor data, hospital, specialist, stock quantity, etc.





Pseudocode:

TIJER || ISSN 2349-9249 || © December 2024, Volume 11, Issue 12 || www.tijer.org

OUTPUT:

6. DRAWBACKS

Data Quality: The model will be abstinent with the practicum data; however, incomplete or noisy data discharges the accuracy of the model and randoms in labelling during the practicum do not appear to have a trend in real time.

Model Limitations: Random Forest models suffer from overfitting on small datasets, random forest, however, hyperparameter tuning is limited hence, the performance of the model.

No Real-Time Updates: A static-based model does not keep pace with dynamic fluctuations of doctor-time schedules and medicines running out.

TIJER || ISSN 2349-9249 || © December 2024, Volume 11, Issue 12 || www.tijer.org

User Experience: Console-based output does not offer user-friendly interfaces and visualization tools for decision-making.

Ethics implications: Data privacy risks and the biases of biases in the data result in equipoise inequities in predictions.

7. CONCLUSION

Predictive analysis of medicines and doctors in government hospitals is a very transformative approach with great implications. Using data-driven insights, predictive models can help enhance medicine supply chains through predicting medicine demand based on historical data, seasonal trends, and patient demographics. This reduces the instances of stockouts and overstocking to ensure that patients have access to drugs whenever needed.

Improve Doctor Availability: Using scheduling data and patient influx trends, predictive analytics can optimize resource allocation, with enough doctors available during peak times, but not so many during lowinflux periods.

Assist in Proactive Health Management: Predictive tools will identify patterns in patient admissions, predict surges in demand, and enable hospitals to strategically allocate resources, thereby reducing wait times and improving care delivery.

Reduce Costs and Inefficiencies: Precise predictions will assist hospitals to optimize procurement and workforce to cut unnecessary wastages that reduce operational efficiencies.

Build Strong Public Health Outcomes: Closing gaps regarding medicine availability and doctors on the ground actually leads to care that is well improved while cutting wait time and assuring that treatments can be timely served.

In conclusion, the integration of predictive analytics in government hospitals is a step toward smarter healthcare management. It not only resolves logistical challenges but also strengthens the trust and efficiency of public healthcare systems. However, successful implementation requires robust data collection systems, trained personnel, and supportive policies to maximize its benefits.

8.REFERENCE

1. Shaikh Karnool Afsa, Khandhakar Nayeem Rejwam, Dada Khalandar "Predictive mechanism for medicines availability in government health centers", publications on ijeat ,2024

2.M.D Boomija,M.I Almas Banu,k Anu priya "Analysis on medicine and Doctor in government hospital", Publications on ijraset, 2019

3.Predictive Analytics on healthcare, www.ijsr.net

4.Smitha Jhajharia,Seema verma,Manish kumar "Predictive Analytics for better health and disease reduction", publication on predictive analytics today ,2021

5.Damien S. E Brokharst,Rogier van de wetering,Ward Ooms,Remko W.Helms "Deploying Predictive Analytics to enhance patient agility and patient value in hospital", publications on Elsevier ,2023

6. Sharique Ahmad, Priyesh Srivastava Tanish Baqar "Predictive Analytics in medical healthcare: a metaAnalysis", publication on Research gate, 2024

7.Mohammed Badawy,Nagy Ramadan and Hesham Ahmed hefny "Healthcare predictive analytics using ML and DL technique: a survey", publications on Springer open,2023

TIJER2412097 TIJER – INTERNATIONAL RESEARCH JOURNAL www.tijer.org a776

TIJER || ISSN 2349-9249 || © December 2024, Volume 11, Issue 12 || www.tijer.org

8.Wullinarllur Raghupathi and Viju Raghupathi "Big data analytics in. Healthcare: promise and potential", Publications on Springer 2014

9.Yichung Wang, Leeann Kung, Choachi Ting "Beyond a technical perspective: understanding big data capabilities in healthcare", publication on ResearchGate,2021

10. Neeraj ,Pradeep kumar "Predictive analysis on availability of medicines and doctors in government hospitals", publications on ijeat,2020

11. Guanhua Lee, Sijia Wang,Lian Leng "Leveraging on Predictive Analytics to Manage Clinic No-Show and Improve Accessibility of Care", publications on IEEE,2017

12. Mimoh Oiha,Dr Kirti Mathur "Proposed Application of Big Data Analytics in Healthcare at Maharaja Yashwantrao Hospital", Publications on IEEE,2016

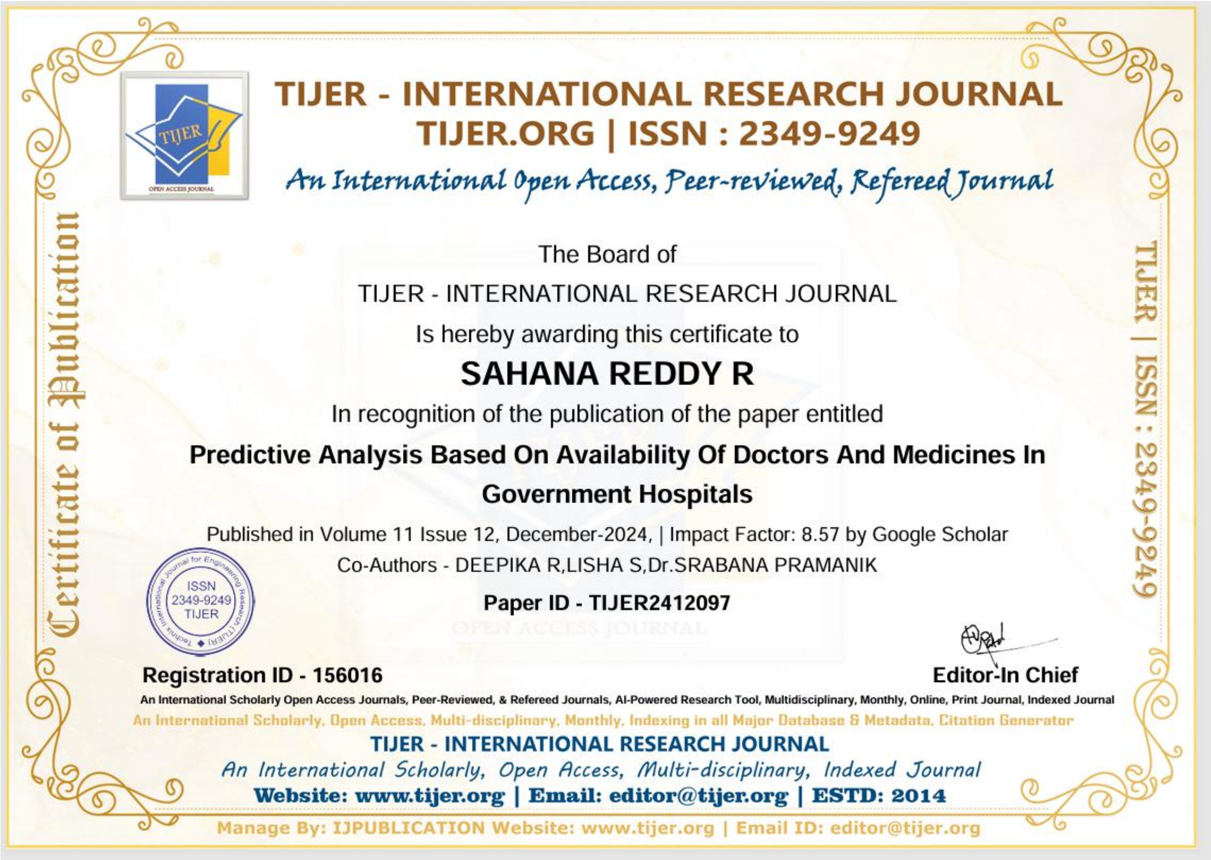
PREDICTIVE ANALYSIS ON AVAILABILITY OF DOCTORS AND MEDICINES IN GOVERNMENT HOSPITALS

ORIGINALITY REPORT

7%	5%	3%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.ijraset.com Internet Source	2%
2	Submitted to University of Adelaide Student Paper	<1%
3	mmcalumni.ca Internet Source	<1%
4	Ashwini Tuppad, Shantala Devi Patil. "Data Pre-processing Issues in Medical Data Classification", 2023 International Conference on Network, Multimedia and Information Technology (NMITCON), 2023 Publication	<1%
5	Xinyuan Song, HSIEH,WEI-CHE, Ziqian Bi, Chuanqi Jiang, Junyu Liu, Benji Peng, Sen Zhang, Xuanhe Pan, Jiawei Xu, Jinlang Wang. "A Comprehensive Guide to Explainable AI: From Classical Models to LLMs", Open Science Framework, 2024 Publication	<1%



TIJER - INTERNATIONAL RESEARCH JOURNAL
TIJER.ORG | ISSN : 2349-9249
An International Open Access, Peer-reviewed, Refereed Journal

Certificate of Publication

TIJER | ISSN : 2349-9249

The Board of
TIJER - INTERNATIONAL RESEARCH JOURNAL
Is hereby awarding this certificate to

SAHANA REDDY R

In recognition of the publication of the paper entitled

Predictive Analysis Based On Availability Of Doctors And Medicines In Government Hospitals

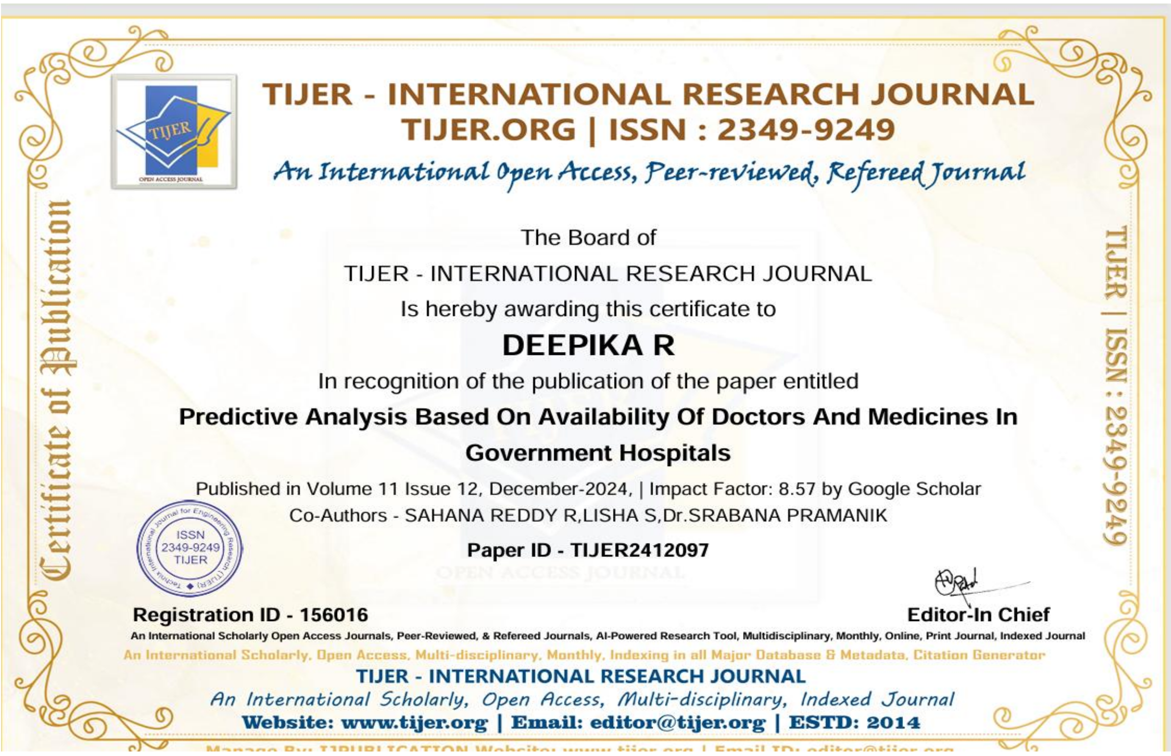
Published in Volume 11 Issue 12, December-2024, | Impact Factor: 8.57 by Google Scholar
Co-Authors - DEEPIKA R,LISHA S,Dr.SRABANA PRAMANIK

Paper ID - TIJER2412097

ISSN 2349-9249 TIJER

Registration ID - 156016 **Editor-In Chief**

An International Scholarly Open Access Journals, Peer-Reviewed, & Refereed Journals, AI-Powered Research Tool, Multidisciplinary, Monthly, Online, Print Journal, Indexed Journal
An International Scholarly, Open Access, Multi-disciplinary, Monthly, Indexing in all Major Database & Metadata, Citation Generator
TIJER - INTERNATIONAL RESEARCH JOURNAL
An International Scholarly, Open Access, Multi-disciplinary, Indexed Journal
Website: www.tijer.org | Email: editor@tijer.org | ESTD: 2014
Manage By: IJPUBLICATION Website: www.tijer.org | Email ID: editor@tijer.org



TIJER - INTERNATIONAL RESEARCH JOURNAL
TIJER.ORG | ISSN : 2349-9249
An International Open Access, Peer-reviewed, Refereed Journal

Certificate of Publication

TIJER | ISSN : 2349-9249

The Board of
TIJER - INTERNATIONAL RESEARCH JOURNAL
Is hereby awarding this certificate to

DEEPIKA R

In recognition of the publication of the paper entitled

Predictive Analysis Based On Availability Of Doctors And Medicines In Government Hospitals

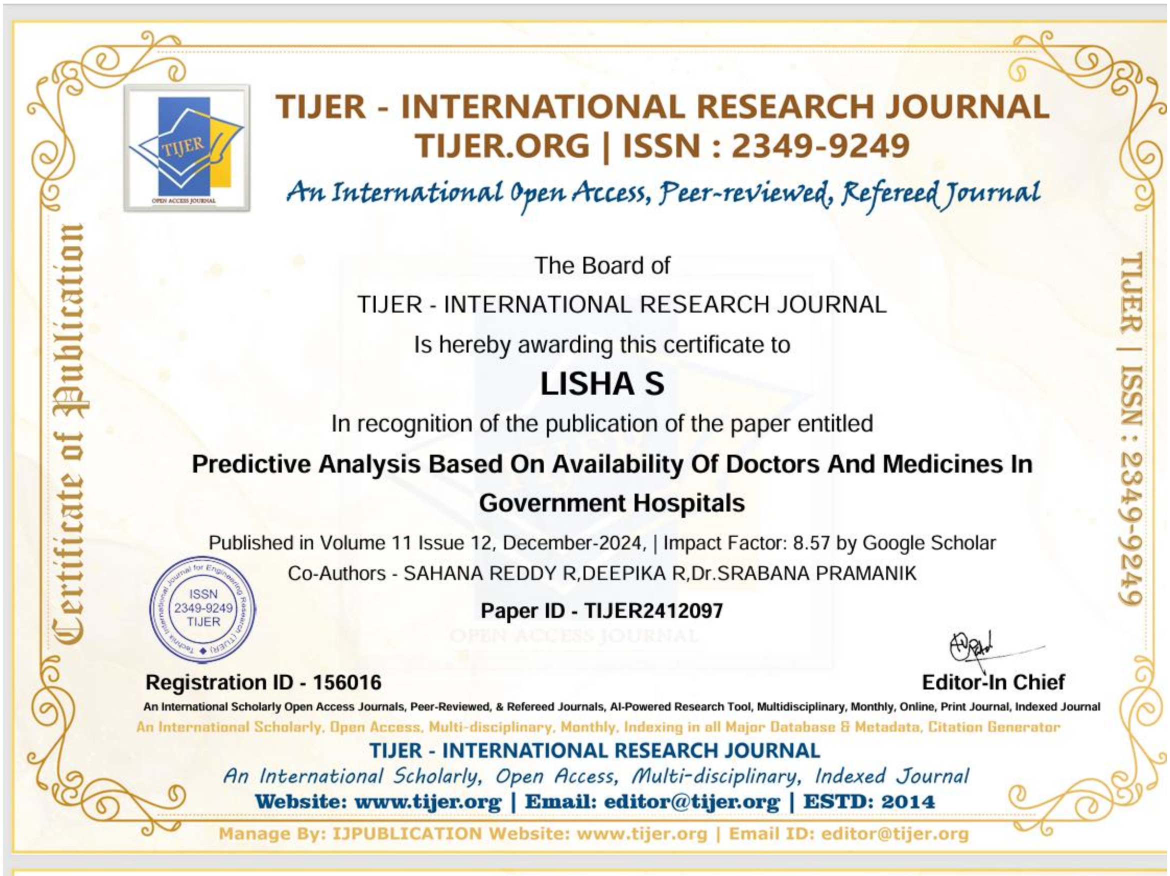
Published in Volume 11 Issue 12, December-2024, | Impact Factor: 8.57 by Google Scholar
Co-Authors - SAHANA REDDY R,LISHA S,Dr.SRABANA PRAMANIK

Paper ID - TIJER2412097

ISSN 2349-9249 TIJER

Registration ID - 156016 **Editor-In Chief**

An International Scholarly Open Access Journals, Peer-Reviewed, & Refereed Journals, AI-Powered Research Tool, Multidisciplinary, Monthly, Online, Print Journal, Indexed Journal
An International Scholarly, Open Access, Multi-disciplinary, Monthly, Indexing in all Major Database & Metadata, Citation Generator
TIJER - INTERNATIONAL RESEARCH JOURNAL
An International Scholarly, Open Access, Multi-disciplinary, Indexed Journal
Website: www.tijer.org | Email: editor@tijer.org | ESTD: 2014



Certificate of Publication

TIJER - INTERNATIONAL RESEARCH JOURNAL
TIJER.ORG | ISSN : 2349-9249
An International Open Access, Peer-reviewed, Refereed Journal

The Board of
TIJER - INTERNATIONAL RESEARCH JOURNAL
Is hereby awarding this certificate to

LISHA S

In recognition of the publication of the paper entitled

Predictive Analysis Based On Availability Of Doctors And Medicines In Government Hospitals

Published in Volume 11 Issue 12, December-2024, | Impact Factor: 8.57 by Google Scholar
Co-Authors - SAHANA REDDY R,DEEPIKA R,Dr.SRABANA PRAMANIK

Paper ID - TIJER2412097

ISSN 2349-9249 TIJER

Registration ID - 156016

Editor-In Chief

TIJER | ISSN : 2349-9249

An International Scholarly Open Access Journals, Peer-Reviewed, & Refereed Journals, AI-Powered Research Tool, Multidisciplinary, Monthly, Online, Print Journal, Indexed Journal
An International Scholarly, Open Access, Multi-disciplinary, Monthly, Indexing in all Major Database & Metadata, Citation Generator
TIJER - INTERNATIONAL RESEARCH JOURNAL
An International Scholarly, Open Access, Multi-disciplinary, Indexed Journal
Website: www.tijer.org | Email: editor@tijer.org | ESTD: 2014
Manage By: IJPUBLICATION Website: www.tijer.org | Email ID: editor@tijer.org

Sustainable Development Goals (SDG)

SDG 3: Good Health and Well-being

Predictive analytics directly improves public health outcomes by ensuring availability of medicines and healthcare professionals, reduced wait times, and timely treatments. Such brings in better health and well-being for all.

SDG 10: Reduced Inequalities

Predictive models help bridge gaps in healthcare access, especially in underserved populations in remote or low-income areas, by providing equitable distribution of medicines and health services.

SDG 17: Partnerships for the Goals

Implementation of successful predictive analytics thus requires the synergy of governments, technology providers, healthcare institutions, and policymakers. Shared knowledge and innovative ideas lead towards sustainable development through such collaboration.