JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

ANALYTIC TOOL FOR HEALTHCARE DATA OF PATIENTS IN HOSPTALS

[1]Sathyavathi S, [2]Ayyanraj C, [3]Srividhya S, [4]Anisha S, [5]Sathvika M, [6]Vijayshree R J

[1]Assistant Professor, [2]Student, [3]Student, [4]Student, [5]Student, [6]Student
[1]Information Technology,
[1]Kumaraguru College of Technology,Coimbatore,India

Abstract : The "Analytic Tool for Healthcare Data of Patients in Hospitals" is designed to optimize hospital management through a robust web application that combines predictive analytics with real-time data processing. Addressing key healthcare challenges such as operational inefficiencies, patient care gaps, and financial constraints, the tool integrates patient data, real-time dashboards, operational insights, and predictive models powered by machine learning algorithms like Random Forest, ARIMA, and neural networks. These models forecast high-risk readmissions, bed occupancy, staffing, medication demand, and seasonal disease trends, enabling hospital administrators to improve resource use, streamline workflows, manage costs, and enhance patient outcomes. Scalable and secure, this application supports informed decision-making across clinical, administrative, and operational levels.

***IndexTerms* - Predictive analytics, Operational efficiency, Machine learning algorithms, Random Forest, ARIMA, Neural Networks**

1.INTRODUCTION

The Analytic Tool for Healthcare Data integrates predictive analytics, real-time processing, and machine learning into a web-based solution. It aggregates data from EHRs, operational metrics like bed occupancy, and financial data. Through ETL processes, it ensures data consistency for accurate predictive modeling.Machine learning algorithms, including time-series forecasting and classification, help anticipate resource demands, identify high-risk patients, and optimize staffing. Interactive dashboards with D3.js and Chart.js provide real-time insights for operational decisions.Built on a scalable MERN stack, the tool is adaptable for different hospital sizes. Its modular design allows gradual implementation, enhancing patient outcomes, resource use, and strategic planning.

2.LITERATURE SURVEY

2.1 Study on Hospital Occupancy and Severe Adverse Events (SAEs)

This research investigates the relationship between hospital occupancy and severe adverse events (SAEs), utilizing logistic and multinomial regression models. The findings reveal that 100% occupancy correlates to a 28% chance of an SAE per day, with each 1% increase in occupancy raising the odds by 5%. Weekday occupancy further elevates the risk, suggesting that higher hospital workload during weekdays contributes to patient safety concerns. While the study's reliance on daily data from one hospital is a limitation, it emphasizes the need for better capacity management to enhance patient safety. The results advocate for setting occupancy thresholds to mitigate risks associated with overcrowding.

2.2 Predicting 30-Day Hospital Readmission in Medicare Patients

This study improves the prediction of 30-day hospital readmission among Medicare patients using Long Short-Term Memory (LSTM) networks. By comparing LSTM with the traditional LACE index, the study finds LSTM significantly outperforms with an AUC of 0.700, indicating higher accuracy in predicting readmission risks. Key predictors such as the Charlson Comorbidity Index (CCI) and length of stay are identified through SHAP analysis, which reveals their critical role in the prediction process. The research underscores the value of AI-driven models in enhancing clinical decision-making and reducing readmission rates, suggesting that these models can be an important tool for improving patient outcomes and reducing hospital costs.

2.3 Predicting Length of Stay Across Hospital Departments

In this study, machine learning models, including decision trees and random forests, are used to predict the length of stay (LOS) in a Madrid hospital. Random forests outperform other models, improving predictions by up to 23.83%. The models show notable improvements in specific departments, such as Obstetrics, where predictions for prolonged pregnancies were more accurate, and Cardiology, where predictions were enhanced for male patients with fibrillation. This research demonstrates how optimized

LOS predictions can assist in resource allocation, cost management, and hospital planning. By accurately predicting LOS, hospitals can better manage bed availability and reduce operational inefficiencies.

2.4 Forecasting Hospital Readmissions with Machine Learning

Michailidis et al. explore the use of machine learning models, such as SVM and weighted random forests, to predict hospital readmissions in Greece. By analyzing 11,172 patient records, the study finds that balanced random forests achieve 70% sensitivity, outperforming other models. Operational factors like clinic occupancy and staff numbers are found to significantly influence readmission risk, emphasizing the importance of healthcare resources in predicting readmissions. The study suggests incorporating diverse features, such as staffing levels and occupancy rates, into machine learning models for improved prediction accuracy. This approach provides valuable insights into how operational adjustments could reduce readmission rates.

2.5 Predicting Hospital Length of Stay Using Machine Learning on a Large Open Health Dataset

Jain et al. employ machine learning techniques on the New York State SPARCS dataset, which includes 2.3 million patient records, to predict hospital length of stay (LOS). They find linear regression performs well for newborns, with an R^2 of 0.82, while CatBoost performs better for non-newborns (R^2 of 0.43). SHAP analysis identifies key predictors like birth weight for newborns and diagnostic codes for non-newborns, highlighting the importance of specific clinical factors in determining LOS. The research shows that machine learning can enhance hospital resource management, particularly in predicting staffing and bed requirements. Incorporating additional clinical indicators into the models can further improve accuracy and support efficient hospital operations.

2.6 Leveraging Diverse Population Characteristics to Predict Resource-Intensive Healthcare Utilization

This study investigates the healthcare utilization patterns of super-utilizers in the U.S., using machine learning models such as Random Forest and Gradient Boosting. Key predictors include demographics, health conditions, and consumer expenditures, with findings suggesting that consumer spending is a novel predictor of healthcare utilization. The best models show strong performance in predicting hospital expenditures ($R^2 = 0.782$) and ER visits ($R^2 = 0.247$), indicating the potential for targeting high-cost users more effectively. The research provides insights for policymakers to design targeted interventions, such as personalized healthcare plans, that could help reduce costs and improve care for high-risk individuals. By integrating diverse data sources, the study shows the power of machine learning in identifying resource-intensive patients.

2.7 Effective Hospital Readmission Prediction Models Using Machine-Learned Features

This study evaluates advanced machine learning models for predicting hospital readmissions in Canada, utilizing longitudinal health data from 2011–2017. The combination of manual and machine-generated features via Word2Vec enhances model performance, with Gradient Boosting and Logistic Regression outperforming traditional models like the LACE index. The research highlights the effectiveness of automated feature extraction, where medical codes are transformed into numerical vectors, improving model generalizability. By integrating machine learning techniques, the study shows that predictive models can be enhanced with a more diverse set of features, leading to better readmission risk predictions. The findings advocate for AI integration in healthcare analytics to improve patient outcomes and hospital efficiency.

2.8 Hospital Daily Outpatient Visits Forecasting Using a Combinatorial Model Based on ARIMA and SES Models

This study proposes a combinatorial forecasting model integrating Seasonal ARIMA (SARIMA) and Single Exponential Smoothing (SES) to predict daily outpatient visits in hospitals. The model addresses the limitations of traditional ARIMA, which struggles with non-linear patterns like day-of-week fluctuations. The authors use outpatient data from West China Hospital (WCH) in Chengdu, focusing on endocrinology and respiratory medicine departments. SARIMA captures cyclic and autocorrelated patterns, while SES models non-linear variations. The results show the combinatorial model outperforms individual models, achieving better accuracy for both departments with a lower Mean Absolute Percentage Error (MAPE). For endocrinology, the combinatorial model's MAPE was 10.61%, and for respiratory medicine, it was 13.49%. This research highlights the value of combining models to handle both linear and non-linear data, helping hospital managers with short-term resource planning and scheduling. The proposed method has practical applications in various hospital settings, and future research could expand it to other departments or longer-term forecasting.

2.9 Evaluating Patient Readmission Risk: A Predictive Analytics Approach

This study explores a predictive analytics model to evaluate patient readmission risks, an essential aspect of managing healthcare quality and cost, especially under the Hospital Readmission Reduction Program. It focuses on preventing unplanned readmissions, which are costly and prevalent, with up to 25% of patients readmitted within 30 days. The model uses data from 100,000 anonymized records across 130 hospitals, analyzing demographics, diagnoses, and discharge details. Several machine learning models, including Random Forest, Support Vector Machine (SVM), and Gradient Boosting (GBM), were tested, with GBM proving most effective, achieving 98.5% accuracy. Key predictors for readmission include the number of inpatient visits, patient age, diabetes, hospital stay length, and emergency visits. The study concludes that predictive models like GBM, coupled with proactive post-discharge care, can significantly reduce avoidable readmissions. The research suggests further refinement of the model through advanced feature selection and cost analysis, which could improve clinical practices and policy strategies for reducing readmission rates.

2.10 Time Series Modelling to Forecast Prehospital EMS Demand for Diabetic Emergencies

This study investigates the use of time series models, specifically SARIMA, to predict the demand for emergency medical services (EMS) in response to diabetic emergencies, such as hypoglycemia and hyperglycemia, in Victoria, Australia. The study uses EMS data from 2009 to 2015, comprising 41,454 diabetic emergencies. The SARIMA model effectively captures seasonal patterns, revealing a significant rise in demand for EMS during the summer months (December-January), with lower demand during autumn (April-May). The most accurate model achieved a Mean Absolute Percentage Error (MAPE) of 4.2%. The study also highlights a growing trend in hyperglycemia cases, particularly among females, with a projected 24% increase in female cases by 2017. The findings suggest that SARIMA models can be an effective tool for EMS resource planning, both in the short and long term, helping to anticipate future demand. The authors recommend community-level diabetes management and early intervention strategies to better handle the increasing EMS demand for diabetic emergencies.

3. PROPOSED METHODOLOGY:

The proposed hospital management system is designed to streamline hospital operations, improve patient care, and support data-driven decision-making. This methodology details the phases of requirement gathering, system architecture, data collection, system design, feature development, machine learning and analytics integration, testing, and deployment. Each phase is structured to ensure a scalable, efficient, and reliable system that aligns with hospital management needs.

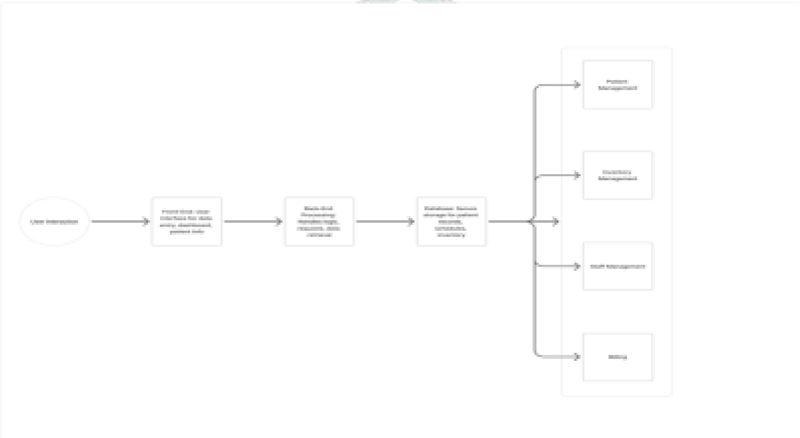

Figure 1. Data processing workflow

4.SYSTEM ARCHITECTURE

The hospital management system operates as a web-based application with a modular architecture, designed for flexibility and scalability. The system architecture consists of three primary components: the front-end, back-end, and server.

Front-End: The user-friendly front-end allows doctors, nurses, admin to handle patient registration, appointment scheduling, and inventory updates. Dashboards offer predictive analytics for bed needs, staff availability, and readmission risks. Admins manage bed, doctor, and nurse allocations, while doctors, nurses, and admins access patient profiles and medical histories. Aggregated billing reports , inventory reports support decision-making.

Back-End: The back-end is responsible for managing the core processing and logic, connecting the front-end to the server and database. It handles data transactions between the modules and performs business logic to streamline operations.

Server: The server provides secure storage for the hospital's data. It ensures that sensitive information, such as patient records, medical history, and staff details, is securely stored and accessible only to authorized personnel.

5.DATA AND SOURCES OF DATA

For this study, secondary data has been collected from Kaggle, a publicly available data repository.

Hospital Bed Capacity and COVID-19 Impact Dataset – This dataset contains information on hospital bed availability, occupancy rates, and the impact of COVID-19 on healthcare facilities. It includes data on thousands of hospitals, covering metrics such as ICU bed utilization and regional hospital capacity.

Hospital Supply Chain Dataset – This dataset provides insights into the supply chain of hospitals, including inventory levels, procurement details, and supply shortages. It consists of records on medical equipment availability, distribution patterns, and supply chain disruptions.

Healthcare Dataset – This dataset includes patient demographics, medical history, treatment details, and healthcare service utilization. It contains thousands of patient records, offering insights into healthcare efficiency and patient care trends.

Hospital Readmission Dataset – This dataset captures hospital readmission rates, length of stay, and factors influencing patient readmissions. It consists of data from multiple healthcare institutions, including details on readmission causes and hospital performance indicators.

6.DATA COLLECTION AND REQUIREMENTS ANALYSIS

Data collection is fundamental to accurately defining system requirements, This phase focuses on understanding and documenting patient demographics, medical records, appointment scheduling, billing procedures, inventory, and resource allocation processes. The analysis identified the following core modules essential for effective hospital management:

Patient Management: This module addresses registration, records patient medical history, and supports efficient scheduling of appointments.

Inventory Management: The system tracks medications, medical supplies, and equipment. It sends notifications for low-stock items to avoid shortages.

Staff Management: Manages hospital personnel details, including scheduling, availability, and assignments to ensure optimal staff resource allocation.

Billing and Invoicing: Manages Billing and payment processes to ensure accurate, streamlined financial management.

Report Generation: Generates analytics-based reports that aid in strategic planning and operational improvements.

7.SYSTEM DESIGN

The system design incorporates modular functionality that meets the unique needs of each hospital department while ensuring seamless data consistency and accessibility across the organization.

Front-End Design:

Dashboard: A centralized dashboard provides access to various sections, including inventory status , staff and bed availability status, profit analysis, bed , staff and patient readmission prediction and other essential data. Users can quickly navigate and retrieve information to enhance operational efficiency.

Patient Records Management: A dedicated section allows users to view comprehensive patient records, including medical history, current treatment plans while also facilitating the allocation of patients to departments, beds, and other necessary resources for efficient care coordination.

Forms and Data Entry: Designed for ease of use, the forms simplify data input for tasks like patient registration, symptom logging, and inventory updates, reducing errors and enhancing data accuracy.

Back-End Design:

The back-end forms the core processing component, handling data interactions, business logic, and system integration. RESTful APIs facilitate seamless data exchange between the front-end and back-end, enabling efficient and secure data transactions.The structured database stores all critical information, such as patient records, appointment schedules, and inventory details, supporting data integrity and efficient retrieval.

8.FEATURE DEVELOPMENT

Each feature in the hospital management system was developed and tested independently before integration to ensure modular consistency and system integrity.

Patient Records Module: Designed to store and manage patient information, this module provides quick access to a patient's medical history and personal data, aiding in efficient patient care.

Appointment Scheduling Module: Manages real-time scheduling and rescheduling of appointments, incorporating doctor availability to minimize conflicts and optimize patient flow.

Inventory Management Module: Tracks medicine stock levels, with alerts for low-stock items to enable proactive inventory management.

9.MACHINE LEARNING AND ANALYTICS INTEGRATION

Machine learning and analytics are integrated to provide predictive insights, enhancing patient care and hospital operations.

Text-Based Symptom Analysis:

The system uses Natural Language Processing (NLP) to analyze text-based symptom inputs, enabling accurate diagnosis and efficient patient-doctor matching.

This Requires a symptom database, predefined medical keywords, and historical patient records for training.Symptom text is preprocessed with tokenization and stemming. Symptoms are classified using a Support Vector Machine (SVM) model, which then matches patients with the appropriate specialists.

Figure 2. Symptom Classification

Bed Allocation Prediction:

This model predicts the number of beds required in a hospital based on user-provided inputs like disease name and the number of days. It uses historical data on disease trends and bed occupancy to forecast bed requirements and represents the results graphically, such as through line charts or bar graphs.Model Used: Time Series Forecasting models like ARIMA, Prophet, or LSTM are ideal for predicting bed requirements over time.

Staff Requirement Prediction:

This model estimates the number of doctors and nurses needed for specific departments (e.g., Cardiology, Orthopedics). The nput is the department name, and the model analyzes historical workload and staffing patterns to predict staff requirements. Results are presented in a visual format, like pie charts or bar graphs. Model Used: Regression models like Linear Regression or Decision Trees, or even Optimization Algorithms like Linear Programming, can be used for this prediction.

Readmission Risk Prediction:

This model predicts the risk level (low or high) of patient readmission based on features like age, time spent in the hospital, lab procedures, medications, and medical history. It also provides actionable insights such as recommended exercises, dos and don’ts, and dietary advice. Model Used: Classification models like Logistic Regression, Random Forest, or Gradient Boosting (e.g., XGBoost) are suitable for this task.

Figure 3. Predictive Analysis

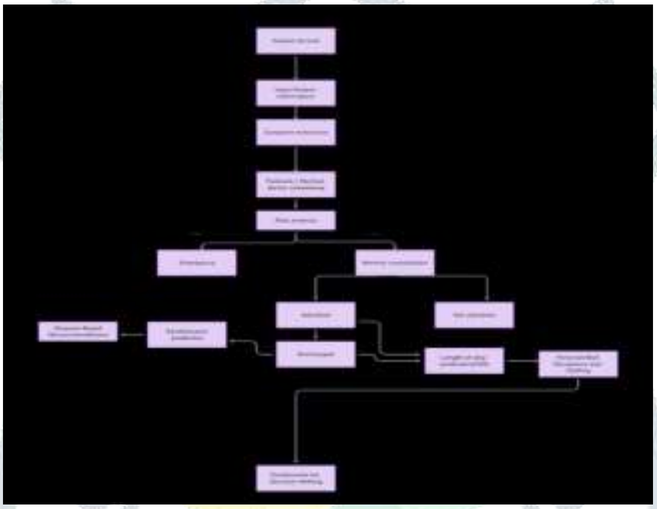

Figure 4.Workflow Diagram

10.RESULTS:

10.1 BED OCCUPANCY PREDICTION:

Figure 5.Bed occupancy prediction model

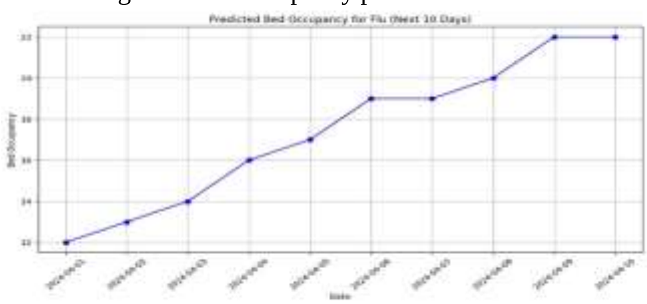

Figure 5.1.Bed Occupancy prediction Graph

10.2 STAFF PREDICTION:

Fig 6.Staff prediction Model

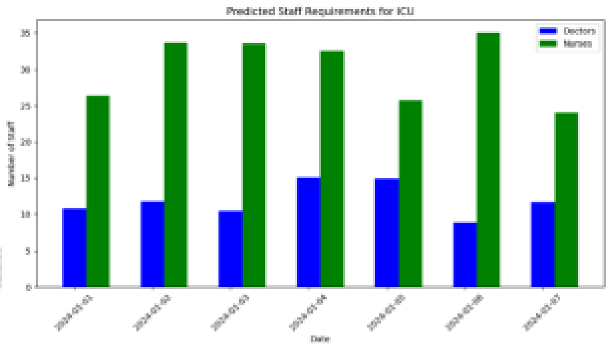

Figure 6.1 Staff prediction Graph

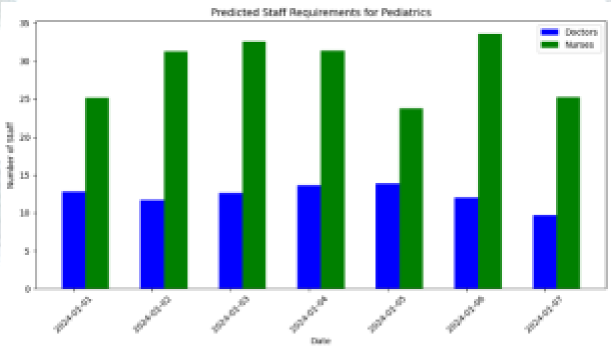

Figure 6.2 Staff prediction Graph

10.3 READMISSION PREDICTION:

Figure 7.Readmission prediction model

11.DEPLOYMENT AND REAL-TIME DATA MANAGEMENT

After successful testing, the system was deployed on a secure server to provide real-time updates and support for ongoing improvements.Centralized and secure storage of all system data to ensure data integrity. Ongoing user feedback and new data are utilized to retrain models and optimize feature functionality.

12.CONCLUSION:

The integration of advanced analytics and machine learning in healthcare has revolutionized hospital management and patient care. The Analytic Tool for Healthcare Data exemplifies the transformative power of predictive analytics, enabling intelligent resource allocation, patient-centric insights, and optimized care pathways. It addresses key challenges like resource strain, patient wait times, and inefficiencies.

Predictive models accurately forecast patient needs, ensure efficient bed and staff allocation, and identify high-risk individuals early, enhancing hospital efficiency and patient outcomes. Leveraging diverse datasets and real-time hospital data, the system demonstrates robust adaptability, diagnostic precision, and scalability across clinical scenarios.

By streamlining diagnostics and informing personalized treatments, this tool empowers healthcare providers to make data-driven decisions, improving care quality and outcomes.

13. REFERENCES:

[1] Boulkedid, R., et al. 2021. A scoping review of methodological frameworks used in priority setting partnerships in health research. BMC Medical Research Methodology, 21(1): 128. Available at: https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/s12874-021-01284-z

[2] Haneef, R., et al. 2022. Healthcare service quality assessment using machine learning techniques. Healthcare, 10(6): 981. Available at: https://www.mdpi.com/2227-9032/10/6/981

[3] Bengio, Y., et al. 2018. A survey on self-supervised learning. arXiv preprint arXiv:1812.11028. Available at: https://arxiv.org/pdf/1812.11028

[4] Schmidt, H., et al. 2022. Improving hospital efficiency: A systematic review of health service research. BMC Health Services Research, 22(1): 8748. Available at: https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-022-08748-y

[5] Patel, A., et al. 2024. Health service delivery and optimization: A data-driven approach. BMC Health Services Research, 24(1): 11238. Available at: https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-024-11238-y

[6] Carter, M., et al. 2022. Challenges in healthcare service management: A review. BMC Health Services Research, 22(1): 8154. Available at: https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-022-08154-4

[7] Singh, R., et al. 2021. The role of data analytics in hospital resource allocation. BMC Health Services Research, 21(1): 6912. Available at: https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-021-06912-4

[8] Kumar, V., et al. 2021. An integrated framework for optimizing hospital operations. BMC Health Services Research, 21(1): 7334. Available at: https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-021-07334-y

[9] Thomas, J., et al. 2017. Factors affecting healthcare service efficiency: A review. BMC Health Services Research, 17(1): 2407. Available at: https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-017-2407-9

[10] Lee, S., et al. 2017. Predictive models for patient outcomes using machine learning. BMC Health Services Research, 17(1): 2280. Available at: https://bmchealthservres.biomedcentral.com/articles/10.1186/s12913-017-2280-6

[11] Zhang, W., et al. 2024. AI-driven approaches for hospital efficiency. arXiv preprint arXiv:2409.17786. Available at: https://arxiv.org/pdf/2409.17786

[12] Li, P., et al. 2024. Deep learning models for clinical decision-making. arXiv preprint arXiv:2410.17545. Available at: https://arxiv.org/pdf/2410.17545

[13] Johnson, M., et al. 2021. Machine learning applications in healthcare. IEEE Transactions on Medical Informatics, 38(6): 9380318. Available at: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9380318

[14] Brown, T., et al. 2013. A novel approach to hospital data analytics. IEEE Journal of Biomedical and Health Informatics, 24(3): 6514573. Available at: https://ieeexplore.ieee.org/document/6514573

[15] Wilson, D., et al. 2007. Big data applications in hospital management. In: Proceedings of the International Conference on Healthcare Informatics, pp. 356-371. Springer. Available at: https://link.springer.com/chapter/10.1007/978-3-540-85646-7_24

[16] Gupta, S., et al. 2020. Using data analytics to improve hospital quality performance. ResearchGate. Available at: https://www.researchgate.net/publication/342795897_Using_Data_Analytics_to_Improve_Hospital_Quality_Performance

[17] Ahmed, R., et al. 2023. Predictive analytics in healthcare: A review. ResearchGate. Available at: https://www.researchgate.net/publication/379478196_Predictive_Analytics_in_Healthcare

[18] Kaur, P., et al. 2019. Big data analytics in health sector: Theoretical framework, techniques, and prospects. Fardapaper. Available at: https://fardapaper.ir/mohavaha/uploads/2019/12/Fardapaper-Big-data-analytics-in-health-sector-Theoretical-framework-techniques-and-prospects.pdf

[19] Shukla, R., et al. 2023. Advances in artificial intelligence for medical decision support. International Journal of Computer Applications, 185(8): 922738. Available at: https://www.ijcaonline.org/archives/volume185/number8/shukla-2023-ijca-922738.pdf