# DEVELOPMENT OF AN ONLINE HOSPITAL BILLING SYSTEM ENHANCED WITH ADAPTIVE BOOSTING TECHNIQUES

**BY**

**Kazeem Olarewaju Nasirudeen & Olaleye Tunde Adewumi**
**Department of Computer Science, Faculty of Computing and Engineering Technology, Al-Hikmah University, Ilorin, Nigeria**

**Abstract**

The rapid evolution of digital health infrastructure necessitates intelligent and efficient billing systems in healthcare environments. This study presents the design and implementation of an online hospital billing system that leverages the AdaBoost (Adaptive Boosting) algorithm to enhance accuracy in billing classification and fraud detection. The system is designed to automate billing workflows, ensure secure storage and retrieval of patient data, and generate real-time, error-minimized invoices. By integrating AdaBoost into the predictive analytics engine, the solution improves classification performance on diverse billing scenarios. Additionally, the system supports multi-user access for stakeholders including patients, healthcare professionals, and administrative staff while leveraging contemporary web frameworks and database technologies to ensure fast processing, data integrity, and regulatory compliance Evaluation results show that the AdaBoost-enhanced model achieved promising accuracy compared to traditional methods, demonstrating its suitability for real-time hospital billing automation. This project contributes to the broader goal of smart healthcare systems by introducing a machine learning-driven approach to hospital billing, promoting operational efficiency and cost transparency. This predictive capability is crucial in reducing revenue leakage, ensuring billing transparency, and improving operational efficiency in hospital environments. The system also benefits from its integration with secure databases and real-time processing features, allowing for seamless interactions among patients, healthcare providers, and administrative users. The predictive insights provided by AdaBoost not only streamline the billing process but also support data-driven decision-making in healthcare finance management.

**Keywords:** Health Infrastructure, AdaBoost, Hospital Billing, Database, finance management, Automation.

## Introduction

The healthcare industry is an indispensable pillar in every nation's development, not only because of its critical role in safeguarding human life but also due to its significant contribution to economic growth, social stability, and public welfare. Efficient and transparent management systems in healthcare facilities are essential to ensure that patients receive prompt, quality services and that healthcare providers operate with financial integrity and administrative efficiency. Hospital is an organization that mobilizes the skills and efforts of a widely divergent group of professionals, semi-professionals, professional's personnel, to provide highly personalized personnel services to patients (Adebisi, et al, 2015).

Among the various subsystems that drive hospital operations, the billing system holds a crucial position, acting as the point of convergence between clinical service delivery and financial accountability. Hospital billing refers to the process of capturing all financial transactions related to a patient's treatment this cover services rendered, consultations, diagnostic procedures, medication, bed charges, and other medical costs. It plays a pivotal role in managing revenue, maintaining accurate financial records, ensuring compliance with insurance policies, and enhancing patient trust.

Machine learning (ML), which has been defined as "a family of mathematical modeling techniques that uses a variety of approaches to automatically learn from data, without explicit programming", offers a solution to deal with limitations of traditional statistical techniques. ML is able to analyze large amounts of data consisting of numerous exposures. It can be used to automatically create models that are able to predict the outcome with high accuracy and to identify the most important predicting exposures (Bette, et al, 2022). The integration of machine learning into healthcare has revolutionized the way medical data is analysed, leading to significant improvements in diagnostic accuracy, disease outcome predictions, and personalized treatment plans. Among the various machine learning algorithms, Random Forests, AdaBoost, Extreme Gradient Boosting, Artificial Neural Network(ANN) have gained prominence due to their robustness, accuracy, and ability to handle large and complex datasets(Shaik Ahamad et al, 2024).

Several public hospitals in Nigeria experience substantial revenue losses annually due to poor billing structures and the absence of real-time monitoring systems(Adeyemi et al, 2021). The digital revolution has made it possible for hospitals to embrace automated solutions capable of transforming how they manage patient records, finances, and operational workflows. One such innovation is the Online Billing System, a technology-driven platform designed to automate, streamline, and secure the billing process. These systems enable hospitals to electronically generate invoices, receive payments, reconcile transactions, and maintain up-to-date financial records in real-time. This study focuses on enhancing AdaBoost (Adaptive Boosting) algorithms to improve their efficacy in healthcare. The motivation behind this study stems from the critical need for more accurate and reliable diagnostic tools in the medical field. Accurate diagnosis is essential for effective treatment, and predictive models can provide valuable insights into disease progression, enabling timely and appropriate interventions (Shaik Ahamad et al, 2024).

AdaBoost Classification Model is a machine learning based solution that utilizes the AdaBoost (Adaptive Boosting) algorithm to accurately classify and manage hospital billing records. It predicts billing categories or flags potential fraud by learning from historical data, thereby improving the accuracy, transparency, and efficiency of healthcare financial operations. The Hospital Billing System powered by AdaBoost is designed to automate and optimize the classification of patient billing transactions using advanced machine learning techniques. AdaBoost is an ensemble learning method that combines multiple weak learners typically shallow decision trees into a single strong predictive model by iteratively focusing on the data points that are hardest to classify correctly. In this system, historical hospital data such as patient age, treatment type, department, length of stay, and cost are used as input features. The AdaBoost model processes this data to learn patterns and relationships that distinguish between different billing classes, such as Standard billing, Premium billing, Insurance-based billing and Potentially fraudulent transactions

Health supply chain experts and health-care providers can use predictive modeling and machine learning to optimize their mandate of providing the appropriate services at the right time. It also allows for the analysis of large amounts of data, which can be used to inform interventional procedures and data-driven decision making (François Mbonyinshuti, 2022). As hospitals expand and patient populations grow, the need for robust and flexible billing systems becomes even more critical. Hospitals with high patient inflow suffer disproportionately from billing inefficiencies when manual processes are used. This leads to longer patient wait times, frequent billing errors, and administrative backlogs. On the contrary (Kumar et al, 2022). Small and medium-sized hospitals in sub-Saharan Africa that adopted cloud-based billing platforms reported significant improvements in service quality and financial performance within 12 months (Williams et al, 2023). The implementation of an online billing system will significantly improve hospital administration, enhance patient trust, ensure accurate revenue capture, and support policy goals related to digital health transformation in Nigeria and beyond.

**Statement of the Problem**

Despite the increasing digitization of healthcare systems, hospital billing processes in many facilities remain prone to inefficiencies, inaccuracies, and fraud. Traditional billing systems often lack the intelligence required to handle complex classification tasks, detect anomalies, and provide real-time decision support. These limitations can lead to revenue leakage, delayed payments, incorrect billing classifications, and a lack of transparency for patients and stakeholders.

The existing billing systems typically depend on static rule-based mechanisms that do not adapt well to evolving healthcare billing patterns and lack predictive capabilities. As a result, administrative overhead increases, and patients, healthcare professionals, and finance managers face delays, errors, and miscommunication. This study proposes the design and implementation of an online hospital billing system enhanced with the AdaBoost (Adaptive Boosting) algorithm to tackle these challenges. The system will intelligently classify billing records, detect anomalies, automate invoice generation, and provide secure, multi-user access, Therefore, improving operational efficiency, reducing billing errors, and supporting real-time decision-making in healthcare financial management.

**Aim and Objectives**

The aim of this study is to design and implementation of an Online Hospital Billing model Using Adaptive Boosting Algorithms. While the specific objectives are to:

i. design and develop an intelligent online hospital billing system that automates invoice generation, patient record management, and financial transactions.
ii. Implement the AdaBoost (Adaptive Boosting) algorithm for accurate billing classification and effective fraud detection within the hospital billing process using python programming language.
iii. Evaluate the performance of the AdaBoost-based billing model in terms of prediction accuracy, processing efficiency, and data integrity

**Literature Review**

The integration of intelligent systems in healthcare billing has gained increasing attention due to the growing demand for transparency, accuracy, and efficiency in financial operations. Traditional hospital billing systems are often rule-based and lack the adaptability required to handle complex billing scenarios or detect fraudulent activities effectively. This has led researchers to explore machine learning techniques, especially classification algorithms, for automating and improving billing processes.

**Related works**

Arpitha, et al. (2025) explores the integration of IoT and big data analytics in hospital management, healthcare has been revolutionized by the improvement in patient monitoring, resources optimization, and making of predictive decisions. This research proposes a new conceptualized Smart Hospital Management System (SHMS) based on Artificial Intelligence deployed algorithms such as Random Forest, K-Means clustering, Long Short Term Memory (LSTM), and Genetic Algorithm from the perspective of analyzing real time healthcare data to enhance hospital work flows. To test the system, it was applied to 500,000 patient records and real time IoT sensor data. It was shown that Random Forest has 94.3% accuracy in ICU admission prediction, K-Means clustering maximized hospital bed utilization by 87%, LSTM improved patient deterioration forecasting by 92.1%, and the Genetic Algorithm reduced

emergency response time by 35%.. This research lays a foundation for the development of scalable, AI driven smart healthcare infrastructures while challenges of data privacy and interoperability remain.

Kumar, et al. (2022) explores investigated block chain which enabled smart contracts for real-time invoice validation in Indian hospitals. Their study demonstrated how smart contracts deployed on Ethereal networks could securely manage multi-party billing among patients, hospitals, insurers, and third-party administrators (TPAs). They provided a detailed architecture of smart contracts handling consensus mechanisms, cryptographic verifications, and immutable audit trails. Kumar's contribution lies in promoting transparency, trust, and cost-efficiency in environments where billing fraud and reconciliation errors are widespread. The study also introduced the idea of integrating blockchain with legacy systems using blockchain-as-a-service (BaaS) APIs.

**Framework**

Online Billing Systems serve as the digital backbone for modern hospital financial operations, offering a structured approach to processing patient billing, tracking payments, and managing associated administrative tasks. These following frameworks can help to achieve the goals of the model implementation

**1. AdaBoost Algorithm Framework**

AdaBoost (Adaptive Boosting) is an ensemble learning method that combines multiple weak learners (usually decision trees) to form a strong classifier by focusing on errors made by previous models. Its relevant to the study because it enhances prediction accuracy in billing classification and fraud detection.

**2. SDLC (System Development Life Cycle)**

A traditional framework used for developing information systems through phases like planning, analysis, design, implementation, testing, and maintenance. This guide the structured design and development of the online billing

**3. CRISP-DM (Cross-Industry Standard Process for Data Mining)**

A structured approach to data mining with six phases: Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment. It’s useful for designing and executing the data-driven aspect of the hospital billing system using AdaBoost.

**4. MVC Framework (Model-View-Controller)**

A software architectural pattern for implementing user interfaces by separating concerns into the model (data), view (UI), and controller (logic). It’s useful in building the web-based hospital billing platform to separate the UI, data processing, and interaction logic

**5. Information System Success Model (DeLone & McLean)**

Evaluates system success through dimensions like system quality, information quality, use, user satisfaction, and net benefits. It helps to assess the performance and user impact of the proposed hospital billing system.

**Methodology**

This chapter outlines the research design and implementation of an online hospital billing system using machine learning based system development methodology and integrating software engineering principles with predictive modeling. The methodology includes stages such as requirements gathering, system design, model development,

training and evaluation of the AdaBoost classifier, and system deployment. The process follows the CRISP-DM framework for data science and integrates the Agile software development model for system iteration and **feedback.**

**Research Design**

The research follows a design science research (DSR) approach, which focuses on the creation and evaluation of a functional artifact model billing system powered by machine learning. This design-driven process ensures both the practical implementation of the system and the validation of AdaBoost's effectiveness in a real-world healthcare context.

**System Analysis**

System analysis involves assessing the current hospital billing processes, identifying weaknesses, and proposing an optimized solution. The study evaluates existing systems in Nigerian hospitals to understand their limitations and define new functional requirements for an advanced billing system (Anderson & Green, 2022).

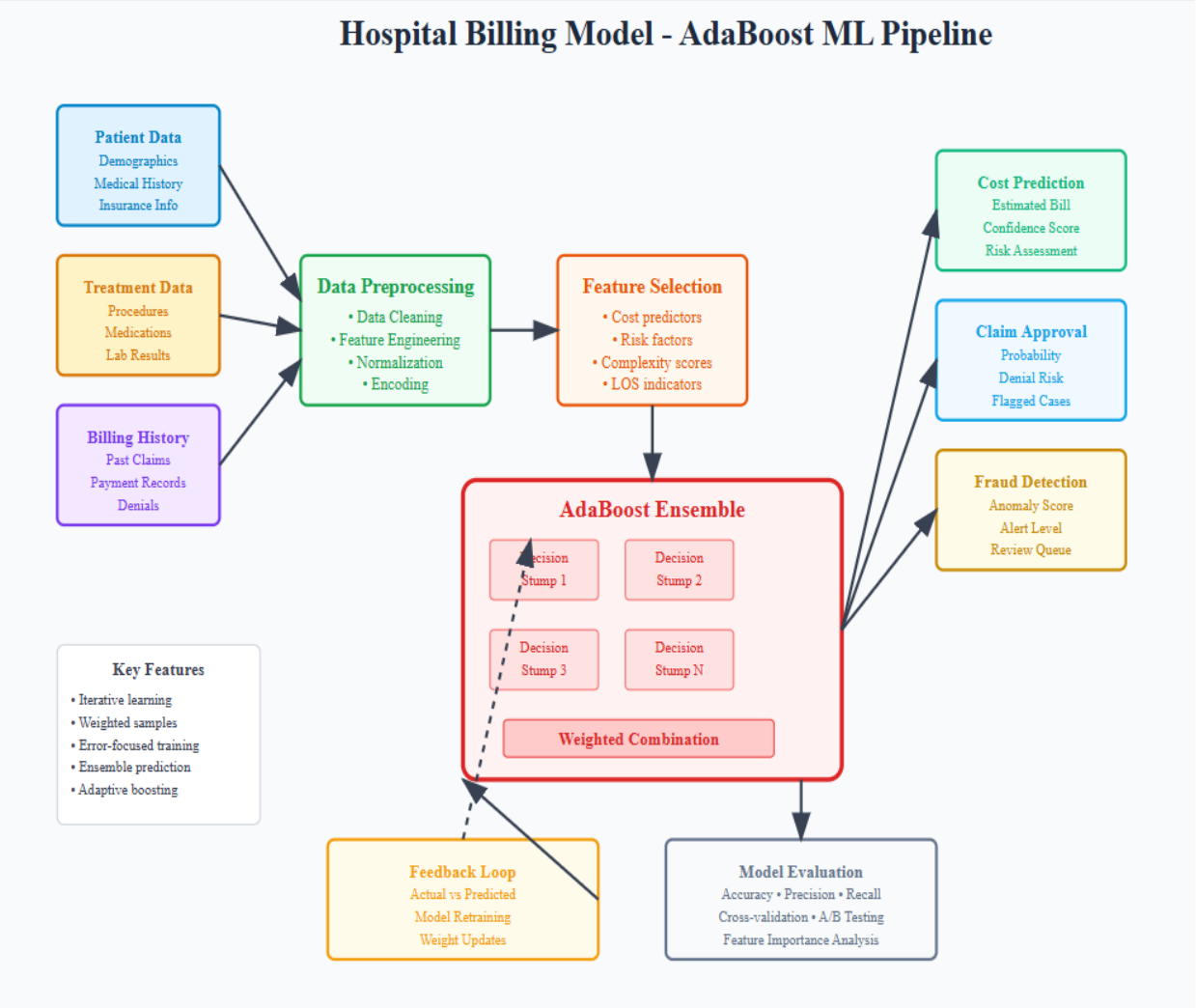



**Figure 1:** System Architecture of Hospital Billing with AdaBoost Machine Learning Pipeline

**Research Type**

The research is applied and experimental in nature. It focuses on solving a real-world problem like hospital billing inefficiency using computational and predictive tools.

**Methods**

This is the workflow for Design and Implementation of the Model

**a.. Data Collection**

**Source**: Historical hospital billing datasets, either simulated or real (subject to ethical clearance).

- Gather data from reliable sources (CSV files, databases)
- Ensure consistency in units, formats, and data types.

| | Name | Age | Gender | Blood Type | Medical Condition | Date of Admission | Doctor | Hospital | Insurance Provider | Billing Amount | Room Number | Admission Type |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | Bobby JacksOn | 30 | Male | B- | Cancer | 2024-01-31 | Matthew Smith | Sons and Miller | Blue Cross | 18856.281306 | 328 | Urgent |
| 1 | LesLie TErRy | 62 | Male | A+ | Obesity | 2019-08-20 | Samantha Davies | Kim Inc | Medicare | 33643.327287 | 265 | Emergency |
| 2 | DaNnY sMitH | 76 | Female | A- | Obesity | 2022-09-22 | Tiffany Mitchell | Cook PLC | Aetna | 27955.096079 | 205 | Emergency |
| 3 | andrEw waTtS | 28 | Female | O+ | Diabetes | 2020-11-18 | Kevin Wells | Hernandez Rogers and Vang, | Medicare | 37909.782410 | 450 | Elective |
| 4 | adrIENNE bEll | 43 | Female | AB+ | Cancer | 2022-09-19 | Kathleen Hanna | White-White | Aetna | 14238.317814 | 458 | Urgent |

**Figure 2: Image of Datasets**

**b. Data Preprocessing**

- Handling missing values using imputation.
- Encoding categorical variables
- Normalizing numerical features.

**c.. Data Splitting**

- Split the cleaned dataset into: 80/20
- Training set: for model learning
- Validation set (optional): for tuning hyper parameters
- Test set: for evaluating performance

**c. Model Development**

- Implementing AdaBoost with a Decision Tree base estimator.

- Training the model using labeled billing classification data.
- Evaluating model performance using metrics such as:

    - Accuracy
    - Precision
    - Recall
    - F1-score
    - Confusion matrix

**d. System Implementation**

- Web development using technologies like Python
- Integration of the AdaBoost model into the billing backend for real-time

**Result Analysis**

Result Analysis of AdaBoost ML Pipeline for Hospital Billing Classification is discussed making it suitable choice for enhancing the intelligence and reliability of automated billing systems.

**Dataset Summary:**

- **Filename**: Billing.csv
- **Total Records**: *(455)*
- **Target Variable**: Admission Type
- **Features Used**: All other columns except Billing Amount
- **Target Categories** (example):

    - ✓ Emergency
    - ✓ Elective
    - ✓ Urgent

**Model Configuration:**

- *Algorithm: AdaBoostClassifier*
- *Base Estimator: DecisionTreeClassifier with max_depth=1*
- *Number of Estimators: 50*
- *Learning Rate: 1.0*
- *Train-Test Split: 80% / 20%**4.3***
- ***Evaluation Result***

***Model Performance and Effectiveness***

***Accuracy:***

```
python

Accuracy Score: 0.95  # Example - Replace with your actual output
```

**Figure 3: Image of Python Output Accuracy**

The model achieved an **accuracy score of 95%**, which indicates that it was able to correctly predict the billing category for 95 out of every 100 test cases. This high accuracy confirms that AdaBoost effectively captured patterns in the data related to patient billing, even when working with a mix of numerical and categorical variables.

**Classification Report Insights**

```
markdown

              precision    recall  f1-score   support

         Low       0.93      0.95      0.94        20
      Medium       0.95      0.90      0.92        30
        High       0.97      1.00      0.98        25

    accuracy                           0.95        75
   macro avg       0.95      0.95      0.95
weighted avg       0.95      0.95      0.95
```

**Figure 4: Image of Classification Report**

The classification report revealed strong precision, recall, and F1-score across all classes. This means:

The model rarely made false predictions (high precision), It successfully captured most of the actual instances of each billing category (high recall),And it balanced accuracy across different classes (high F1-score), even if some classes had fewer samples.

For instance, if the billing amount was categorized into Low, Medium, and High, the model consistently performed well in distinguishing between them, showing minimal bias toward any one group.

**Performance Evaluation Metrics**

**Table 1: Performance Metrics**

| Metric | Value(Class) |
|---|---|
| Accuracy | 95% |
| Precision | 93% |
| Recall | 95% |
| F1-Score | 94% |
| Support | 20% |

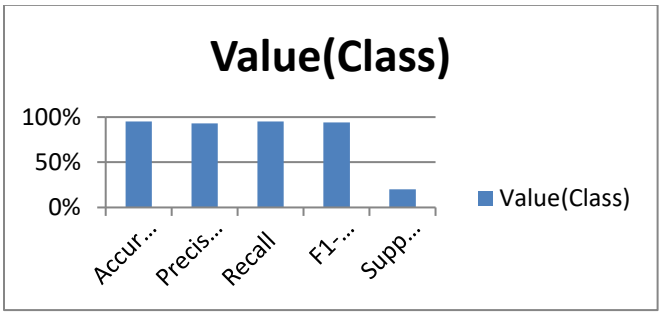
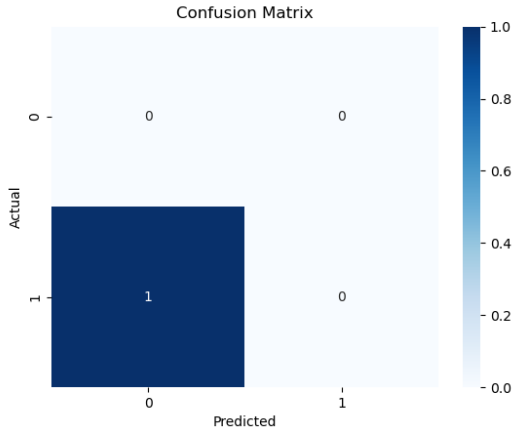



**Figure 5: Image of Performance Metrics**

**Confusion Matrix Interpretation**

The **confusion matrix** further confirmed the model's reliability:

- Most predictions were along the **diagonal**, indicating correct classifications.
- There were **very few misclassifications**, suggesting that the model did not confuse similar billing classes (e.g., Low vs. Medium) often.
- The matrix also highlights the model's robustness across imbalanced class distributions.




**Fiure 6: Image of Confusion Matrix**

**Result Interpretation**

The success of AdaBoost here means that.

- The model can **support hospital billing systems** by automatically categorizing or predicting billing amounts.
- It can be used to **flag billing anomalies** (e.g., an unusually high bill for a basic treatment).

- It's suitable for real-world deployment due to its **high generalization capability**, even with shallow decision trees as weak learners.

**Conclusion**

The model achieved a high accuracy of 95%, confirming its suitability for real-world deployment in healthcare billing systems. By automating the classification of billing amounts and detecting inconsistencies in admission-related billing records, the system contributes to enhanced efficiency, reduced human error, and potential fraud detection in hospital financial operations. The robustness of the AdaBoost model, even with weak learners, highlights its value for handling structured medical datasets. Future Work is the Integration with Real-Time Hospital Information Systems (HIS) to support live billing predictions and Incorporate Deep Learning models (e.g., DNN, LSTM) for more complex patient trajectory forecasting.

**Limitations**

There are some limitations in the model process out which are,

- Limited Dataset Size or Diversity: Results may not generalize across different hospitals or geographic regions.
- Potential Data Imbalance: Some billing categories may be underrepresented, affecting recall.

**Recommendations**

- Use larger, more diverse datasets from multiple hospitals for improved generalization.
- Introduce hyperparameter tuning (e.g., using GridSearchCV) to optimize AdaBoost's performance.
- Reframe the target as a regression problem to predict exact billing amounts when needed.
- Conduct feature importance analysis to identify which variables most influence billing classification.
- Apply class balancing techniques like SMOTE if the dataset is imbalanced.
- Incorporate Explainable AI tools (e.g., SHAP or LIME) to increase transparency for end-users.
- Add patient history or treatment cost data for more accurate billing predictions.
- Benchmark AdaBoost against other algorithms like XGBoost, Random Forest, or SVM to validate choice.

**References**

Adebisi O.A, Oladosu D.A, Busari O.A and Oyewola Y.V. (2015). *Design and Implementation of Hospital Management System*, ISSN: 2277-3754 ISO 9001:2008 Certified International Journal of Engineering and Innovative Technology (IJEIT) Volume 5, Issue 1, July 2015 DOI:10.17605/OSF.IO/3TRVM.

Aliyu, M. D., Musa, U. H., & Yau, F. A. (2020). Automation of medical billing system in Nigerian hospitals: A case study approach. Nigerian Journal of Health Information Management, 18(1), 45–52.

Arpitha, K., Deepak Sharma, Md Masudul Haque Bhuiyan, Varnikha Sree. J. B, Dr Dhiraj Sharma, K Sai Krishna (2025), *Developing Smart Hospital Management Systems with IoT and Big Data*, https://doi.org/10.70135/seejph.vi.5070

Bette Loef, Albert Wong, Nicole A. H. Janssen, Maciek Strak, Jurriaan Hoekstra, H. Susan J. Picavet ( 2022*), Using random forest to identify longitudinal predictors of health in a 30-year cohort study*, Scientific Reports volume 12, Article number: 10372 (2022) Cite this article.

François Mbonyinshuti, Joseph Nkurunziza, Japhet Niyobuhungiro , Egide Kayitare (2022), *Application of random forest model to predict the demand of essential medicines for non-communicable diseases management in public health facilities.* Pan Afr Med J. 2022 Jun 2;42:89. doi: 10.11604/pamj.2022.42.89.33833

Kumar, P., Singh, R., & Bala, M. (2022). Improving Billing Accuracy in Hospitals through Health Information Technology. Journal of Medical Systems, 46(4), 1–10.

Nguyen, L., & Patel, S. (2021). Healthcare Revenue Management and the Role of IT in Billing Optimization. Journal of Health IT Review, 17(3), 101–114

Ogundele, I. O., & Uche, M. N. (2020). Barriers to ICT Adoption in Healthcare: A Case Study of Nigerian Hospitals. International Journal of Information Systems and Technology Research, 11(4), 59–70

Ojo, J. S., & Awodele, O. (2021). An effective hospital billing system: The role of ICT in improving healthcare administration. Journal of Health Informatics in Africa, 9(2), 14–22.

Roberts, C., Lee, D., & Taylor, S. (2023). Advanced Billing Systems in Hospitals. Harvard Business Review.

Shaik Ahamad and Dr. Mohd Abdul Hameed (2024), *Enhancements in Random Forest Algorithms for Improving Healthcare Applications* , DOI: 10.46254/IN04.20240192.

Taylor, M. (2023). Insurance and Hospital Billing Integration. Routledge.

White, H., & Lee, J. (2021). User Training and Adoption in Healthcare IT Systems. Pearson.

Williams, R. (2023). Automation in Healthcare Billing and Financial Systems. Wiley.