Journal of Information Systems Engineering and Management
2025, 10(25s), 30195
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

Intelligent FD in Human Body Monitoring Using Hybrid Learning Models - Biomoni-FD

S. Nandini[1], Dr. A.R. Arunachalam[2]

[1]Research Scholar, Department of Information Technology, Dr. M.G.R Educational and Research Institute, Maduravoyal, Chennai, Tamilnadu, India.
[2]Professor & Dean, Department of Computer Science, Dr. M.G.R Educational and Research Institute, Maduravoyal, Chennai, Tamilnadu, India.

Email: [1]sn.nandini21@gmail.com, [2]arunachalam.cse@drmgrdu.ac.in

ARTICLE INFO

Received: 20 Dec 2024

Revised: 11 Feb 2025

Accepted: 26 Feb 2025

ABSTRACT

Influenza is a highly contagious respiratory disease and is still a serious threat to public health all over the world. Forecasting techniques help in monitoring seasonal influenza and other influenza-like diseases and also in managing resources appropriately to formulate vaccination strategies and choose appropriate public health measures to reduce the impact of the disease. The aim of this investigation is to forecast the monthly incidence of seasonal flu in Saudi Arabia for the years 2020 and 2021 using the XGBoost model and compare it with ARIMA and SARIMA models. The results show that the XGBoost model has the lowest values MAE, MAE, and RMSE compared to the ARIMA and SARIMA models and the highest value of R-squared (R^2). This study compares the accuracy of the XGBoost model with ARIMA and SARIMA models in providing a forecast of the number of monthly seasonal influenza cases. These results confirm the notion that the XGBoost model has a higher accuracy of prediction than that of the ARIMA and SARIMA models, mainly due to its capacity to capture complex nonlinear relationships. Therefore, the XGBoost model could predict monthly occurrences of seasonal influenza cases in Saudi Arabia.

Keywords: Monitoring System, FD, Data Preprocessing, Min-Max Normalization, Z-Score Standardization, Decimal Scaling, Isolation Forest, Anomaly Detection, False Alarm Rate, Detection Delay, Machine Learning, Health Monitoring, Real-Time Classification.

INTRODUCTION

With the swift progress in healthcare and wearable technology, HBMS have become indispensable for real-time health assessment, disease prediction, and remote patient observing. These systems utilize sensors and IoT-enabled devices to measure key physiological parameters, including heart rate, blood pressure, body temperature, and oxygen saturation. However, maintaining their accuracy and reliability is vital, as system faults or malfunctions could result in incorrect diagnoses and pose risks to patient safety.

1.1 Significance of IOT Sensors in Healthcare

The integration of Internet of Things (IoT) technologies in healthcare marks a significant transformation, offering numerous practical benefits (Chunyan Li et al., 2024). Sensors act as the core elements of any healthcare system, making their reliability crucial. Ideally, these sensors should be compact, silent, highly accurate, energy-efficient, and capable of minimizing data transmission delays while maintaining optimal performance. Wearable sensors, in particular, must balance precision and size, posing a notable challenge. However, their outputs must also be sufficiently accurate to assist doctors in making informed medical decisions. In contrast, medical-grade sensors, while highly precise, tend to be bulky, difficult to transport, and require specialized equipment along with trained personnel for operation (Mamdiwar, S.D. et al., 2021). The incorporation of IoT sensors into healthcare has revolutionized patient monitoring, diagnostics, and treatment by facilitating real-time data collection and analysis. These sensors, embedded in wearable devices, smart medical equipment, and implantable systems, continuously track vital signs such as heart rate, blood pressure, oxygen saturation, and body temperature. By transmitting this data to cloud-based healthcare platforms, IoT enables remote monitoring, early disease detection, and timely medical intervention, ultimately easing hospital workloads and enhancing patient care. For chronic disease management, IoT-powered devices assist both patients and healthcare providers in tracking conditions such as diabetes, hypertension, and cardiac disorders, allowing for proactive treatment adjustments. Beyond patient monitoring, IoT sensors are instrumental in smart hospitals and medical automation, streamlining workflows, tracking medical equipment, and ensuring drug safety. Integrated sensors within hospital infrastructure help regulate air quality, temperature, and humidity in critical care units, creating optimal conditions for patient recovery. Additionally, IoT-enabled smart pill dispensers improve medication adherence by reminding patients to take their prescriptions on time. By leveraging predictive analytics and AI-driven insights, IoT sensors enhance healthcare efficiency, lower operational costs, and significantly improve patient consequences. In the healthcare sector, IoT technologies are actively enhancing patient care and outcomes by enabling remote monitoring (H.H. Alshammari, 2023; B.G. Mohammed et al., 2023), facilitating personalized treatment plans (S. Tiwari et al., 2023), and optimizing healthcare delivery (S. Krishnamoorthy et al., 2023).

1.2 Importance of ML and DL in HBMS

With the growing adoption of ML techniques, data-driven deep learning (DL) is gaining significant attention in patient care (Chiranjib Chakraborty et al., 2024). FD in HBMS is a crucial aspect of healthcare technology, designed to identify and address system errors before they affect patient care. These faults can result from sensor degradation, signal interference, hardware failures, or software malfunctions. While traditional FD methods rely on threshold-based monitoring, modern approaches leverage ML, artificial intelligence (AI), and DL techniques to enhance fault prediction and classification accuracy. The integration of intelligent computing techniques into HBMS has greatly enhanced health monitoring, FD, and disease prediction. These advanced models process large volumes of physiological data collected from wearable sensors and medical devices, enabling real-time anomaly detection and health risk assessment. By efficiently refining data, they help reduce noise, extract meaningful patterns, and improve overall system accuracy. This, in turn, facilitates early diagnosis, personalized treatment planning, and proactive healthcare interventions, reducing medical emergencies and enhancing patient outcomes. Beyond detecting irregularities in physiological signals, advanced computing methods excel at identifying intricate health patterns that traditional techniques may overlook. They improve the analysis of medical images, continuous health data, and biometric signals, ensuring precise and reliable assessments. Additionally, these intelligent systems adapt to individual health trends, supporting personalized monitoring and automated decision-making. By utilizing vast datasets and self-learning mechanisms, they empower healthcare professionals with more accurate diagnostics, efficient monitoring, and timely interventions, ultimately improving patient care and reducing healthcare costs. This study examines different FD mechanisms in HBMS, comparing traditional methods with intelligent techniques to improve accuracy, efficiency, and reliability. The objective is to design robust FD algorithms capable of autonomously identifying system anomalies, minimizing false alarms, and enhancing healthcare outcomes.

RELATED WORKS

Remote healthcare monitoring remains a critical area of research for both industry and academia. With the advent of Wireless Body Area Networks (WBANs), patient supervision has become more feasible through implanted body sensors that communicate via wireless interfaces. These sensors, despite their compact size and limited resources—such as power, computing, and communication capabilities—are susceptible to faults and potential damage. Therefore, implementing an efficient system to detect faults or anomalies in sensed data is essential. Mohamed Bahache et al., 2022 propose an innovative, optimized, and hybrid approach that combines machine learning and statistical techniques to detect faults in WBANs without compromising device resources or functionality. Experimental results demonstrate that this method achieves a FD accuracy exceeding 99.62%, with a low mean absolute error of 0.61%, significantly outperforming existing state-of-the-art solutions.

The IoT-based healthcare system has reached a peak in popularity due to its vast potential compared to other IoT applications. By integrating sensors with IoT healthcare, patient health data can be effectively collected and analyzed, making IoT-driven healthcare widely accepted. However, several challenges must be addressed to develop a flexible IoT-based healthcare system, including the continuous availability of healthcare professionals and essential facilities in remote areas during emergencies.

Additionally, human-entered data is often less reliable than automatically generated data. The advancement of IoT-based health monitoring systems enables personalized treatment in specific situations, reducing healthcare costs and resource wastage while continuously improving patient outcomes. Mohammad Shahidul Islam et al., 2018 propose an IoT-based system utilizing the MySignals development shield integrated with a low-power, long-range (LoRa) wireless network. The system incorporates various sensors, including an electrocardiogram (ECG) sensor, body temperature sensor, pulse rate sensor, and oxygen saturation sensor, all connected via MySignals and LoRa. The study evaluates sensor performance and wireless platform effectiveness through physiological data analysis and statistical methods. MySignals successfully collects physical health data, transmitting it to a personal computer using a LoRa wireless system. The results confirm that the MySignals platform effectively interfaces with the sensors, successfully implementing communication via the hyper-terminal program, ultimately contributing to the development of an IoT-based healthcare system.

Fitness monitoring systems that utilize sensors, IoT, and cloud servers have gained significant popularity in recent years. In this approach, patients either wear miniature sensors or have them surgically implanted to monitor vital health factors. These sensors collect data and transmit it to a central server for processing and storage, allowing doctors to access the information remotely at their convenience. Since these systems handle critical health data, they must be highly reliable and free from faults. However, various factors, including hardware failures, software issues, and transmission errors, can impact system performance at different levels. To ensure uninterrupted functionality, health monitoring systems should incorporate fault prevention mechanisms that enable seamless data transmission, even in the presence of malfunctioning sensor nodes. KoushikKarmakar et al., 2022 developed a FD and Recovery Framework for sensor-based remote health monitoring, designed to identify faulty sensor nodes and select alternative nodes for continued data transmission. Faulty nodes are prevented from participating in data collection and transmission until they are either repaired or replaced. The study utilizes LibeliumMySignalsHW v2 sensors to collect patient vitals through ECG, SpO2, and temperature sensors. Additionally, an Arduino UNO R3 board functions as the microcontroller device. The acquired sensor data is analyzed using the proposed algorithm and simulated in MATLAB to detect node-level faults. The framework ensures reliable, seamless, accurate, and timely data transmission.

One of the key challenges in Wireless Body Area Networks (WBANs) is the detection of sensor faults. Smrithy Girijakumari Sreekantan Nair et al., 2020 present a method for accurately identifying faulty sensors, which helps distinguish genuine medical conditions from sensor malfunctions, thereby reducing false alarms and enhancing the quality of services provided by WBANs. The proposed Sensor FD (SFD) algorithm is based on Pearson correlation coefficients and simple statistical techniques. It first identifies strongly correlated attributes using Pearson correlation coefficients and then applies the SFD algorithm to detect faulty sensors. The effectiveness of the proposed algorithm was validated using two datasets from the IC (Intensive Care) database, with results compared to those of existing methods. Additionally, the algorithm's time complexity was analyzed relative to other approaches. The findings indicate that the proposed method achieved high detection rates and low false alarm rates, with accuracy levels of 97.23% for Dataset 1 and 93.99% for Dataset 2.

Structural Health Monitoring (SHM) is a non-destructive testing method used to assess the condition and predict the lifespan of civil infrastructure. Sensor faults can lead to the loss of crucial data, inaccurate assessments of structural conditions, and, in the worst case, undetected structural damage. To address this, fault diagnosis (FD) methods have gained traction within the SHM community. However, most traditional FD approaches focus on single faults, which can oversimplify the complexities of real-world SHM systems, where multiple faults may occur simultaneously. Thamer Al-Zuriqat et al., 2023 propose an adaptive FD method for SHM systems that addresses the occurrence of multiple simultaneous faults in sensors. This adaptive approach incorporates FD, isolation, and accommodation, building on analytical redundancy, which uses correlated data from multiple sensors within the SHM system. Specifically, the method detects faults by utilizing artificial neural network (ANN) models, which leverage correlations between sensor data. Once fault occurrence time instances are identified in the sensor data, faults are isolated by analyzing the moving average of individual sensor data around these time points. For fault accommodation, the ANN models are modified by excluding faulty sensors and using pre-fault sensor data to generate virtual outputs that replace the faulty readings. The proposed adaptive FD approach was validated through two tests using sensor data collected from an SHM system installed on a railway bridge. The results show that this approach effectively maintains the accuracy, reliability, and performance of real-world SHM systems, even in the presence of simultaneous faults in multiple sensors.

Accurate and timely fall detection plays a crucial role in healthcare, particularly in monitoring the elderly, where rapid responses can help prevent serious consequences. This study introduces a novel fall detection model based on a transformer architecture, designed to analyze movement speeds of key body points tracked using the MediaPipe library. By continuously monitoring these points in video data, the model detects real-time speed variations that may indicate a fall. The transformer's attention mechanism enables it to capture even subtle changes in movement, leading to an impressive accuracy of 97.6% while significantly reducing false alarms compared to traditional methods. This approach is highly applicable in environments such as elderly care facilities and home monitoring systems, where reliable fall detection ensures timely intervention. By focusing on movement dynamics, the model enhances both accuracy and dependability, making it well-suited for diverse real-world scenarios. Ultimately, it presents a promising solution for improving safety and care for vulnerable populations across various settings.

The IoT plays a vital role in innovative applications across various sectors, including smart cities, smart homes, education, healthcare, transportation, and defense. In healthcare, IoT applications are particularly valuable as they facilitate secure, real-time remote patient monitoring, ultimately enhancing the quality of life. Suliman Abdulmalek et al., 2022 explore emerging trends in healthcare monitoring systems by examining the role of IoT. Their study highlights the significance and advantages of IoT-based healthcare systems, offering a systematic review of recent research on the topic. The literature review compares different systems in terms of effectiveness, efficiency, data protection, privacy, security, and monitoring capabilities. Additionally, the paper examines IoT-based healthcare monitoring systems that utilize wireless and wearable sensors, categorizing different types of healthcare-monitoring sensors. The study also addresses key challenges and unresolved issues related to healthcare security, privacy, and Quality of Service (QoS). Finally, the authors provide recommendations for improving IoT healthcare applications and outline future research directions based on recent technological advancements.

3. Proposed Hbms Framework

Figure 1 presents a structured methodology for FD in a HBMS utilizing the MIMIC-III dataset. It outlines the process of identifying and diagnosing faults within the system, ensuring accurate health monitoring and data reliability.

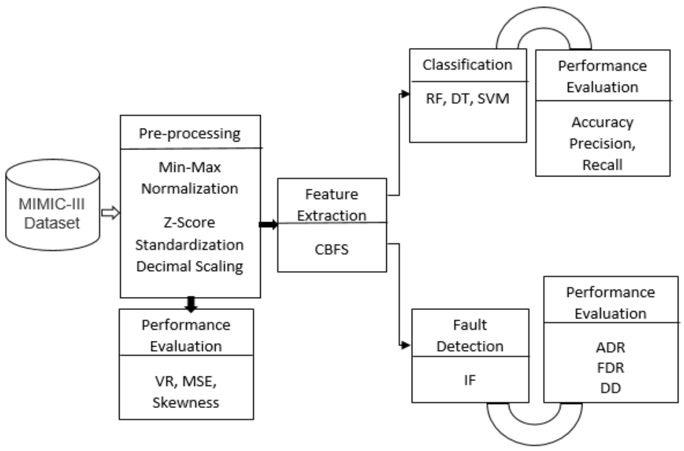

Fig 1. HBMS Framework for FD and Classification

3.1 Data Collection

MIMIC-III (Medical Information Mart for Intensive Care) is a comprehensive, single-center database that contains detailed patient information from critical care units in a major tertiary care hospital.

3.2 Pre-processing

The process starts with data preprocessing, a vital step in normalizing and standardizing raw data before implementing ML techniques. Three preprocessing methods—Min-Max Normalization, Z-Score Standardization, and DS—are applied to transform the dataset. These techniques help maintain consistency in data distribution, ensuring its suitability for accurate analysis.

Min-Max Normalization
Min-Max Normalization is a commonly used preprocessing technique in HBMS that scales physiological and clinical data into a fixed range. In HBMS, raw data from IoT sensors, medical devices, and electronic health records (EHRs) often vary significantly in scale—for example, HR (measured in beats per minute) and blood GL (measured in mg/dL). By applying Min-Max Normalization, these values are standardized while maintaining their relative relationships, ensuring compatibility for FD and classification tasks.

In HBMS, Min-Max Normalization helps to uniformly scale physiological and clinical data, leading to improved model performance in detecting and classifying faults. The normalization process follows the formula:

$$\mathrm{HR}' = \frac{\mathrm{HR} - \mathrm{HR_{min}}}{\mathrm{HR_{max}} - \mathrm{HR_{min}}} - - - (1)$$

$$\mathrm{BP}' = \frac{\mathrm{BP} - \mathrm{BP_{min}}}{\mathrm{BP_{max}} - \mathrm{BP_{min}}} - - - (2)$$

$$\mathrm{GL}' = \frac{\mathrm{GL} - \mathrm{GL_{min}}}{\mathrm{GL_{max}} - \mathrm{GL_{min}}} - - - (3)$$

Where:
- HR', BP', GL' = Normalized values for **Heart Rate (HR), Blood Pressure (BP), and Glucose Level (GL)**
- HR, BP, GL = Original sensor readings
- $\mathrm{HR_{min}}$, $\mathrm{BP_{min}}$, $\mathrm{GL_{min}}$ = Minimum values recorded for respective parameters
- $\mathrm{HR_{max}}$, $\mathrm{BP_{max}}$, $\mathrm{GL_{max}}$ = Maximum values recorded for respective parameters

Z-Score Standardization
This process ensures that all health-related features, including HR, BP, and GL, are scaled consistently, eliminating unit dependencies and enhancing model performance in FD and classification.

$$\mathrm{X}' = \frac{\mathrm{X} - \mu}{\sigma} - - - (4)$$

Where:
- X' = Standardized value
- X= Original sensor reading
- μ= Mean of the dataset
- σ = Standard deviation of the dataset

This transformation enhances FD by making anomalies more distinguishable, as outliers will exhibit notably higher or lower standardized scores. Z-Score Standardization is particularly effective in managing outliers and ensuring that the data follows a normal distribution, making it well-suited for ML models in HBMS.

Decimal Scaling (DS)
DS is a preprocessing technique applied in HBMS to normalize sensor data by scaling values according to the highest power of 10 in the dataset. This method ensures that all physiological parameters, including HR, BP, and GL, are adjusted to a uniform range while preserving the original distribution of values.
The formula for DS is:

$$\mathrm{X}' = \frac{\mathrm{X}}{10^{\mathrm{j}}} - - - (5)$$

DS is especially beneficial for FD and classification models in HBMS, as it preserves the relative magnitude of features while ensuring all values fall within a consistent range. This method is less sensitive to extreme outliers and is straightforward to implement in real-time healthcare applications.

3.3 Feature Selection
After preprocessing, the data undergoes feature extraction using CBFS, which removes redundant attributes while preserving the most significant features. This process is crucial for enhancing computational efficiency and improving model accuracy.

Correlation-Based Feature Selection (CBFS)
CBFS is a crucial method in HBMS that identifies and retains the most significant physiological parameters while discarding redundant or irrelevant features. This approach enhances the performance

of FD and classification models by increasing accuracy, minimizing computational complexity, and reducing the risk of overfitting.

CBFS assesses the relationship between input variables and the target outcome while also examining correlations among input features. The objective is to keep features that strongly correlate with the target variable (e.g., normal vs. abnormal health conditions) while eliminating those that exhibit high inter-correlation, as they contribute redundant information.
The correlation coefficient (**r**) is determined using Pearson's correlation formula:

$$r_{X,Y} = \frac{\sum(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum(X_i - \overline{X})^2 . \sqrt{\sum(Y_i - \overline{Y})^2}}} - - - (6)$$

Where:
- X_i and Y_i are individual feature values
- $\overline{X}$ and $\overline{Y}$ are the mean values of features X and Y
- $r_{X,Y}$ represents the degree of correlation between two variables, ranging from **-1 to 1**
- **r≈1** → Strong positive correlation
- **r≈−1** → Strong negative correlation
- **r≈0** → No correlation

By applying CBFS, excessively similar features are eliminated, ensuring that only unique and meaningful attributes are utilized for IF and classification (RF, SVM, DT) in HBMS. This optimization improves diagnostic precision, enhances computational efficiency, and strengthens real-time monitoring performance in IoT-driven healthcare solutions.

Isolation Forest (IF)

After selecting relevant features, FD is carried out using IF, an anomaly detection method designed to recognize unusual patterns within the dataset. The effectiveness of FD is evaluated through ADR, FDR, and DD, ensuring that abnormal conditions are accurately and promptly identified. The IF-based Algorithm enhances the detection of anomalies, effectively pinpointing irregularities in health data (M. Subramanian et al., 2024).

IF is a highly efficient anomaly detection algorithm employed in HBMS to detect irregular physiological patterns and potential health risks. It is particularly useful for identifying rare but critical health conditions by isolating outliers from normal data distributions.

The fundamental principle behind IF is that anomalies (or outliers) are easier to isolate compared to normal data points. The algorithm constructs multiple random DTs and determines how many splits (partitions) are necessary to isolate a specific data point:

- Normal data points require more splits since they are densely clustered within the dataset.
- Abnormal data points (faults or health anomalies) require fewer splits, as they are rare and significantly different from typical data.

The anomaly score is calculated based on the average path length of a data point across multiple trees. A shorter average path length indicates a higher likelihood of the point being an anomaly.
The anomaly score $s(X)$ is determined using the following formula:

$$s(X) = 2^{-\frac{E(h(X))}{c(n)}} - - - (7)$$

- $h(X)$ = Average path length of instance Xacross all trees
- $E(h(X))$ = Expected path length
- $c(n)$ = Normalization factor for n data points

IF is extensively utilized for detecting health anomalies in real time, enabling early identification of medical conditions and facilitating timely preventive measures.

3.4 Classification

After FD, the dataset moves to the classification stage, where ML models such as RF, DT, and SVM are employed to distinguish between normal and faulty instances.

Random Forest (RF)
RF is a widely used ML algorithm for classification tasks in HBMS**.** It is an ensemble learning technique that constructs multiple decision trees and aggregates their predictions to enhanceaccuracy and reliability. In HBMS, RF plays a vital role in classifying normal and abnormal health conditions by analyzing physiological sensor data.

RF Operates By
1. Creating multiple DTs using different subsets of training data (bootstrapping).
2. Using random FS to reduce overfitting and improve generalization.
3. Aggregating predictions from all trees using majority voting (for classification) to determine the final class label.

The final classification is determined by:

$$\hat{y} = \arg\max_{k} \sum_{i=1}^{N} 1(T_i(X) = k) - - - (8)$$

Where:
- $T_i(X)$ = Prediction of the iii-th tree for input X
- k = Class label (e.g., normal vs. abnormal)
- N = Total number of decision trees

RF is a robust classification algorithm used in HBMS to accurately identify health conditions and potential disease risks. Its capability to handle extensive sensor data with precision makes it highly suitable for real-time IoT-driven healthcare solutions

Decision Tree (DT)
DTis a popular ML technique for classification in HBMS. It classifies normal and abnormal health conditions by applying hierarchical decision rules to sensor data.
How DT Works:
- **Root Node:** Represents the complete dataset.
- **Internal Nodes:** Define decisions based on feature thresholds.
- **Leaf Nodes:** Indicate classification results (e.g., normal or abnormal).
- **Splitting Criteria:** Uses metrics like **Gini Index** or **Information Gain** to enhance classification precision.

The Information Gain (IG) formula for determining optimal splits is:

$$IG = H(\text{parent}) - \sum \frac{|\text{Child}|}{|\text{Parent}|} H(\text{Child}) - - - (9)$$

Entropy, denoted as H(x), quantifies data impurity, where lower values indicate more effective splits. In HBMS, DTs serve as a powerful classification approach, offering transparent and swift predictions for disease detection, risk evaluation, and continuous health tracking. Nevertheless, Random Forest typically yields superior outcomes by decreasing overfitting.

Support Vector Machine (SVM)
SVM is a robust supervised learning technique extensively applied in classification tasks within HBMS. It efficiently differentiates between normal and abnormal health states using sensor data gathered from wearable devices, medical sensors, and IoT-driven health monitoring platforms.
SVM functions by:
1. Transforming data points into a higher-dimensional space to determine an optimal decision boundary.
2. Identifying the hyperplane that maximizes the separation margin between distinct classes (e.g., healthy vs. unhealthy).
3. Utilizing kernel functions (e.g., linear, polynomial, radial basis function) to address complex, non-linear patterns in health-related data.

The SVM decision function is represented as:

$$f(x) = \sum_{i=1}^{N} \alpha_i y_i \, K(x_i, x) + b - - - (10)$$

Where:
- x_i = Training samples
- y_i = Class labels (+1 for normal, -1 for abnormal)
- $K(x_i, x)$ = Kernel function
- α_i = Lagrange multipliers

- b = Bias term

SVM serves as a reliable classification model in HBMS, enabling precise disease diagnosis, anomaly detection, and continuous patient monitoring.

RESULTS AND DISCUSSION

4.1 Dataset Description

MIMIC-III is a comprehensive, single-center database containing detailed records of patients admitted to intensive care units in a major tertiary hospital. It includes vital signs, medications, laboratory results, clinician notes, fluid balance, procedure and diagnostic codes, imaging reports, hospital stay duration, survival outcomes, and other critical patient data.

4.2 Preprocessing Methods Evaluation

Data preprocessing is essential for optimizing FD and classification in HBMS. The efficiency of a preprocessing approach is evaluated using Variance Retention (VR%), Mean Squared Error (MSE), and Skewness. These metrics assess how effectively the method maintains data integrity while promoting consistency and normalization.

VR (%)

Evaluates the extent to which the preprocessing technique retains the original feature variance. Greater values signify superior data preservation post-transformation, with values closer to 100% being ideal.

$$\text{Variance Retention} = \left(\frac{\text{Variance After Preprocessing}}{\text{Variance Before Preprocessing}}\right) \times 100 - - - (11)$$

MSE between Original and Scaled Data

Assesses the level of distortion caused by preprocessing, where lower values signify minimal data loss. A smaller measure indicates better retention of the original data quality, ensuring higher accuracy in subsequent analysis.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(X_{\text{original}} - X_{\text{scaled}})^2 - - - (12)$$

Skewness

Analyses the asymmetry in data distribution both before and after preprocessing. A skewness value of 0 represents a perfectly symmetrical distribution, while values near 0 indicate a well-balanced dataset with minimal distortion.

$$\text{Skewness} = \frac{n}{(n-1)(n-2)}\sum\left(\frac{(X_i - \overline{X})^3}{S}\right) - - - (13)$$

The following Table 1 compares various preprocessing techniques using three critical metrics: VR%, MSE, and Skewness. These measures assess the effectiveness of each method in preserving data integrity and ensuring optimal transformation.

Table 1: Assessment of Preprocessing Techniques on MIMIC-III Dataset

Pre-processing Method	**VR(%)**	**MSE**	**Skewness**
Min-Max Normalization	94.3%	0.005	0.13
Z-Score Standardization	97.2%	0.004	0.11
DS	98.9%	0.001	0.02

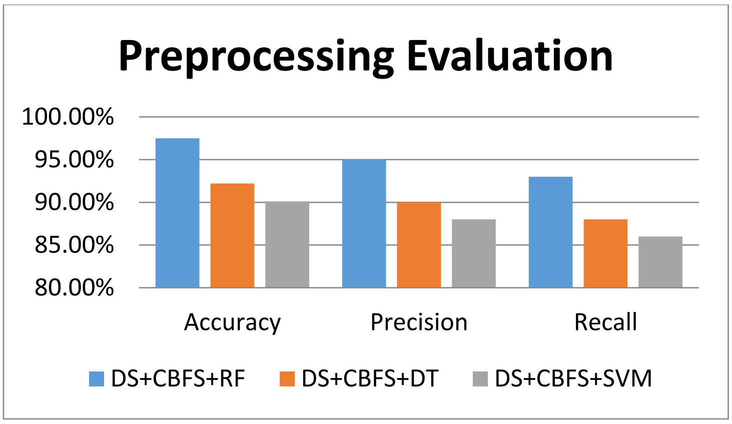

Fig 2. Assessment of Preprocessing Techniques on MIMIC-III Dataset Graph

Variance retention reflects the extent to which the original data variability is conserved following preprocessing. DS achieves the highest variance retention at 98.9%, indicating that it preserves nearly the entire original data distribution. Z-Score Standardization follows closely with 97.2%, making it a strong option for maintaining data structure. In contrast, Min-Max Normalization retains only 94.3%, implying that it may introduce slight distortions, especially in datasets with diverse value ranges. MSE quantifies the transformation error caused by preprocessing. DS yields the lowest MSE (0.001), signifying minimal deviation from the original data. Z-Score Standardization has an MSE of 0.004, reflecting a moderate transformation effect, whereas Min-Max Normalization exhibits the highest MSE (0.005), indicating the greatest alteration in data values. A lower MSE is preferable, as it minimizes information loss and enhances accuracy in subsequent ML applications. Skewness measures the asymmetry of data distribution before and after preprocessing. DS substantially lowers skewness to 0.02, creating a nearly symmetric dataset suitable for models requiring normal distributions. Z-Score Standardization results in a skewness of 0.11, demonstrating moderate improvement in data symmetry. Min-Max Normalization retains the highest skewness (0.13), suggesting that it may be less effective in normalizing distributions, which could impact model performance. Classification models are assessed using Accuracy, Precision, and Recall, which determine prediction reliability. This comprehensive framework facilitates efficient human body parameter monitoring, ensuring precise FD and classification while optimizing computational efficiency.

4.3 Performance Evaluation of IF

Anomaly Detection Rate (ADR)

ADR quantifies the percentage of actual anomalies (faults) accurately detected by the model. Also known as the True Positive Rate (TPR) **or** Sensitivity, it evaluates the model's effectiveness in identifying anomalies within a dataset.

$$\mathrm{ADR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \times 100 - - - (14)$$

Where:

ADR represents the percentage of real anomalies that the model successfully detects. It is commonly referred to as the TPRorSensitivity in anomaly detection. A higher ADR indicates a model's strong capability to identify faults accurately, ensuring reliable anomaly detection in various applications.

A higher ADR, ideally close to 100%, signifies improved FD and accurate anomaly identification. Conversely, a lower ADR leads to more missed anomalies, potentially reducing system reliability and effectiveness.

False Alarm Rate (FPR)

$$\mathrm{FPR} = \left(\frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{FN}}\right) \times 100 - - - (15)$$

A lower FPR results in fewer false alarms, enhancing system reliability and reducing unnecessary alerts.

Detection Delay (DD)

The detection delay is measured as the average time (in seconds) required to identify an anomaly after it occurs.

Table 2: FD Performance using IF Algorithm

Scenario	(ADR) (%)	(FPR) (%)	DD (sec)
Normal Condition	98.5%	2.1%	0.5
Sensor Failure	96.3%	3.4%	1.2
Abnormal Heart Rate	97.8%	2.7%	0.8

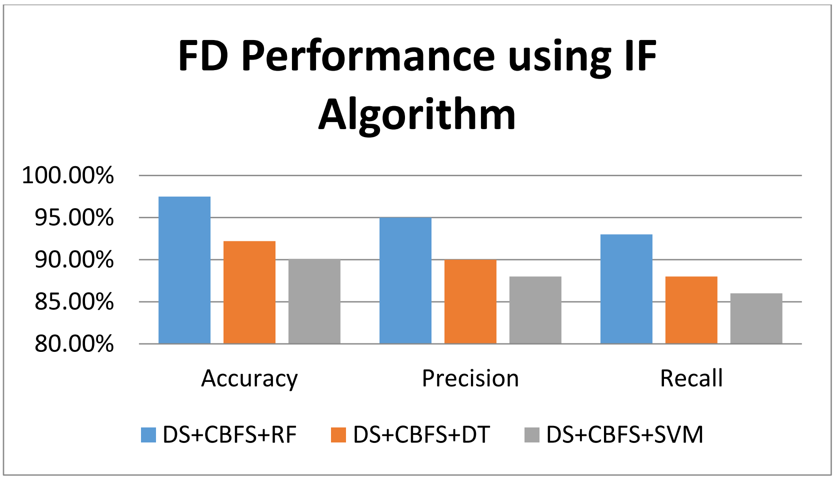

Fig 3. FD Performance using IF Algorithm Graph

4.4 Classification Model Performance

Accuracy

Accuracy evaluates the effectiveness of a classification model in correctly distinguishing between normal and abnormal conditions in HBMS.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} - - - (16)$$

The classification model's performance is determined using the following metrics:

- **TP (True Positives):** The count of actual abnormal conditions accurately identified.
- **TN (True Negatives):** The number of actual normal conditions correctly recognized.
- **FP (False Positives):** Instances where normal conditions are mistakenly classified as abnormal, leading to false alarms.
- **FN (False Negatives):** Cases where the model fails to detect abnormal conditions, resulting in missed anomalies.

Precision

Precision quantifies the percentage of correctly identified abnormal health conditions among all instances classified as abnormal by the model.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} - - - (17)$$

High precision indicates a lower rate of false alarms, ensuring that the model primarily flags actual abnormal conditions.

Recall

Recall assesses the system's ability to identify abnormal health conditions from the total number of actual abnormal cases.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} - - - (18)$$

Table 3 showcases the evaluation metrics for various classification techniques applied inHBMS. The models were assessed using DS for data preprocessing, CBFS for feature extraction, and three distinct classifiers: RF, DT, and SVM.

Table 3: Performance for Various Classification Models

Model Name	Accuracy	Precision	Recall
DS+CBFS+RF	97.5%	0.95	0.93
DS+CBFS+DT	92.2%	0.90	0.88
DS+CBFS+SVM	90.10%	0.88	0.86

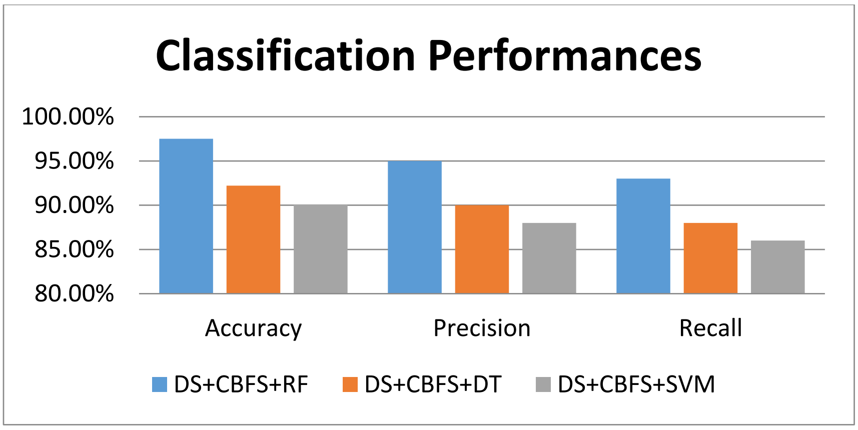

Fig 4. Performance for Various Classification Models Graph

The DS+CBFS+RF model achieved the highest accuracy of 97.5%, demonstrating its ability to correctly distinguish between normal and abnormal conditions. Furthermore, its precision (0.95) and recall (0.93) indicate that 95% of detected abnormal cases were correctly classified, while 93% of actual abnormal instances were successfully identified. This makes RF the most dependable and effective model for HBMS FD, as it significantly reduces both false positives and false negatives. The high recall value ensures that critical health issues are promptly recognized, making RF an excellent choice for practical healthcare applications.

The DS+CBFS+DT model exhibited a slightly lower performance, attaining an accuracy of 92.2%, with a precision of 0.90 and recall of 0.88. While it remains a strong performer, the lower recall implies that 12% of actual abnormal cases were not detected, making it less suitable for applications requiring early identification of critical conditions. However, Decision Trees are computationally simpler than RF, making them an appropriate option for real-time monitoring systems with limited processing resources.

The DS+CBFS+SVM model recorded the lowest accuracy among the three models, achieving 90.10%. It obtained a precision of 0.88 and recall of 0.86, indicating a higher false negative rate compared to RF and DT. This suggests that 14% of actual abnormal cases went undetected, making it a less reliable choice for FD in healthcare, where missing critical conditions could lead to serious consequences. Although SVM is effective in handling complex data distributions, its overall performance in this scenario falls short when compared to RF and DT.

Among all three models, DS+CBFS+RF model stands out as the optimal choice for HBMS classification due to its high accuracy, precision, and recall. It minimizes both false alarms and missed detections, making it the most trustworthy model for real-time monitoring.

CONCLUSION AND FUTURE WORK

HBMS are essential for continuous health assessment and FD. However, the effectiveness and dependability of these systems rely on efficient data preprocessing, feature selection, and classification strategies. This research highlighted the influence of various preprocessing techniques, where Z-Score Standardization, combined with CBFS, delivered superior results by minimizing data variability and improving feature representation. This methodology facilitated stronger model training, leading to enhanced classification performance.For FD, the IF algorithm was employed to detect anomalies within the system, with its efficiency assessed through ADR, FPR, and DD. The findings validated its ability to accurately identify faults, enabling quick responses to irregular conditions. In terms of classification, three ML models—RF, DT, and SVM—were analyzed. Among these, RF achieved the highest accuracy, precision, and recall, establishing it as the most dependable model for HBMS classification.In summary, the proposed framework effectively enhances FD and classification in HBMS, delivering a trustworthy solution for real-time health monitoring. By integrating optimized preprocessing, feature selection, and classification methods, this approach ensures greater accuracy and early fault identification, ultimately improving the safety and effectiveness of health monitoring systems. Future work can investigate DL approaches and real-time sensor data integration to further enhance performance and scalability.

REFERENCES

[1] Mohamed Bahache , Abdou El Karim Tahari, Jorge Herrera-Tapia , NasreddineLagraa , Carlos Tavares Calafate & ChakerAbdelazizKerrache (2022), “Towards an Accurate Faults Detection Approach in Internet of Medical Things Using Advanced Machine Learning Techniques”, Sensors (Basel), 22(15):5893. doi: 10.3390/s22155893

[2] Mohammad Shahidul Islam, Mohammad Tariqul Islam, Ali F. Almutairi, GanKokBeng, NorbahiahMisran&Nowshad Amin (2019), “Monitoring of the Human Body Signal through the Internet of Things (IoT) Based LoRa Wireless Network System”, MDPI, Applied Sciences , Vol. Issue 9, 10.3390/app9091884

[3] Koushik Karmakar, Sohail Saif, Priya Das, Sarmistha Neogy, Suparna Biswas(2022),“DRF: FD and Recovery Framework for Seamless Data Transmission in Remote Health Monitoring Using WBAN”, ACM Digital Library, PeriodicalsWireless Personal Communications: An International Journal, Vol. 123, No. 1. https://registry.opendata.aws/mimiciii/

[4] SmrithyGirijakumariSreekantan Nair, RamadossBalakrishnan(2020), “A precise sensor FD technique using statistical techniques for wireless body area networks“, Wiley Online Library, https://doi.org/10.4218/etrij.2019-0207

[5] Thamer Al-Zuriqat, Carlos ChillónGeck, Kosmas Dragos&Kay Smarsly (2023), “Adaptive Fault Diagnosis for Simultaneous Sensor Faults in Structural Health Monitoring Systems”,MDPI,Infrastructures 2023, 8(3),39; https://doi.org/10.3390/infrastructures803003 9

[6] Duncan Kibet, Min Seop So, Hahyeon Kang, YongsuHan&Jong-Ho Shin(2024), “Sudden Fall Detection of Human Body Using Transformer Model”, MDPI, Sensors 2024, 24(24), 8051; https://doi.org/10.3390/s24248051

[7] SulimanAbdulmalek, Abdul Nasir, Waheb A. Jabbar, Mukarram A. M. Almuhaya, Anupam Kumar Bairagi, Md. Al-MasrurKhan&Seong-HoonKee (2022), “IoT-Based Healthcare-Monitoring System towards Improving Quality of Life: A Review”, Healthcare, Volume 10, Issue 10 , 10.3390/healthcare10101993

[8] Mamdiwar, S.D.; Shakruwala, Z.; Chadha, U.; Srinivasan, K.; Chang, C.-Y. “Recent advances on IoT-assisted wearable sensor systems for healthcare monitoring”. Biosensors 2021, 11, 372.

[9] [Chunyan Li, Jiaji Wang , Shuihua Wang & Yudong Zhang(2024), “A review of IoT applications in healthcare”, Science Direct, Neurocomputing, Vol. 565, 14 January 2024, 127017

[10] H.H. Alshammari (2023), “The internet of things healthcare monitoring system based on MQTT protocol”, Science Direct, Alex. Eng. J., 69 (2023), pp. 275-287.

[11] B.G. Mohammed &D.S. Hasan (2023), “Smart healthcare monitoring system using IoT, Science Direct, iJIM, 17 (01) (2023), p. 141

[12] S. Tiwari, K. Nahak, A. Mishra(2023), “Revolutionizing healthcare: the power of IoT in health monitoring”, Science Direct, J. Data Acquis. Process., 38 (2) (2023), p. 2416

[13] Muhammad Saqib. (2025). Optimizing Spot Instance Reliability and Security Using Cloud-Native Data and Tools. In Journal of Information Systems Engineering and Management (Vol. 10, Issue 14s, pp. 720–731). Science Research Society. https://doi.org/10.52783/jisem.v10i14s.2387

[14] S. Krishnamoorthy, A. Dua, S. Gupta (2023), “Role of emerging technologies in future IoT-driven Healthcare 4.0 technologies: a survey, current challenges and future directions”, J. Ambient Intell. Humaniz. Comput., 14 (1) (2023), pp. 361-407

[15] Chiranjib Chakraborty , Manojit Bhattacharya , Soumen Pal& Sang-Soo Lee (2024), “From machine learning to deep learning: Advances of the recent data-driven paradigm shift in medicine and healthcare”, Current Research in Biotechnology, Vol. 7, 2024, 100164

[16] M. Subramanian, K. Gokulakrishnan, B. V. Nirmala, S. Vijay Shankar, R. S. Krishnan and G. VinothRajkumar, "GuardianForest: Cloud-Powered Patient Health Assurance through Isolation Forests," 2024 International Conference on Inventive Computation Technologies (ICICT), Lalitpur, Nepal, 2024, pp. 2039-2046, doi: 10.1109/ICICT60155.2024.10544839.