Intelligent Information Management, 2025, 17(6), 235-262
https://www.scirp.org/journal/iim
ISSN Online: 2160-5920
ISSN Print: 2160-5912

Analysis of Decision Support Systems (DSS) Integration with AI Using Mobile Health (m-Health) and Wearables

Aisha Alqarni

Department of Information Systems, College of Computer Science and Engineering, Taibah University, Madinah, Saudi Arabia
Email: TU4570421@taibahu.edu.sa

How to cite this paper: Alqarni, A. (2025) Analysis of Decision Support Systems (DSS) Integration with AI Using Mobile Health (m-Health) and Wearables. *Intelligent Information Management*, **17**, 235-262.
https://doi.org/10.4236/iim.2025.176013

Received: September 4, 2025
Accepted: November 10, 2025
Published: November 13, 2025

Abstract

Smart health refers to the integration of cutting-edge technologies into healthcare systems to improve patient care and apply intelligent clinical decision-making. The study investigates how artificial intelligence (AI) can be integrated with decision support systems (DSS) in healthcare using wearable technologies and mobile health (m-Health) platforms. This research focuses on applying artificial intelligence to clinical decision support systems to improve healthcare delivery in real-time. A proposed model aims to implement machine learning algorithms to analyze continuous health data retrieved from wearables and mobile health applications. The research will evaluate the proposed model using two datasets m-Health [1] and Indicators of Heart Disease [2], available online in Kaggle repository. Random Forest, Gradient Boosting Machines, and XGBoost are used to assess the predictive performance within structured sensor data. Only statistically significant variables are retained for feature selection based on their correlation coefficients. Random Forest reports 94.23% accuracy on the m-Health dataset, which affirms the effectiveness in sensor-driven activity performance evaluation. Gradient Boosting achieves 91.26% accuracy on the structured medical dataset, reaffirming its credibility in health risk assessment. It is noted that embedding AI models into decision support systems will act as a framework for responsive, data-driven decisions in medicine. The capability to recognize anomalies and forecast possible health threats facilitates proactive and continual care for patients, thus supporting real-time monitoring. Uniformed data preprocessing on other dataset enhances the experimental validity and reliability. The study demonstrates that AI-empowered CDSS have great capabilities in closing the gaps of delays in diagnosis and the need for timely, tailored interventions. Future research will focus in incorporating real-time wearable data with adaptive AI decision support systems to improve clinical applicability and scalability.

Keywords

Artificial Intelligence (AI), Decision Support Systems (DSS), Wearables, Mobile Health (m-Health), Real-Time Monitoring

1. Introduction

Digital health technology adoption has taken the delivery and management of healthcare services to another dimension [3]. DSS can be used to assist clinicians in making various decisions, and the use of AI has broadened this task. The experience of AI in the field of health data interpretation and acquisition gives way to more effective and rapid decision alternatives [4]. Wearable devices such as mobile health, fitness bands or smart wristwatches track patient vitals, including the heart rate, physical activity, and sleep hours among others [5]. Such data has the potential to be examined almost instantly and can potentially provide a holistic picture of an individual's health status. These new advances have enhanced the early detection of disease, the tracking of even chronic diseases, and the engagement of effective personalized health programs. While AI-enhanced decision support systems in mobile health have shown usefulness [6]. Earlier, researchers focused more on data analytics and early detection attempts. However, many of them are narrowed down to certain diseases rather than having a broader patient-oriented model. The study seeks to fill that gap by examining AI Model for DSS to be predictive and adaptive in communicating with cardiovascular patients. By facilitating data-driven events and actions, the aim is to increase the chances of patients coming to hospitals for preventive care, manage the number of patients who go to the hospital, and reduce the pressure on healthcare systems.

1.1. Research Hypothesis/Null Hypothesis

The key hypothesis of the research states that the integration of AI models in Decision Support Systems (DSS), utilizing real-time physiological data collected via wearable technology, will significantly improve early detection of health risks. It supposes that machine learning algorithms will be able to predict and diagnose health conditions by analyzing behavioral and physiological data, resulting in high accuracy. AI models are expected to determine abnormal trends in heart rate and activity, which are comorbid factors associated with cardiovascular disease. Furthermore, the research hypothesizes that the feature selection techniques can also detect the other critical determinants of health that impact the predictive modeling accuracy. The outcomes are expected to align with medical understanding, as the health determinants predicted by such models are known as risk factors. Whereas the null hypothesis of the research states that AI (ML) models, when applied to wearable data, do not significantly improve predictive accuracy or healthcare decision-making.

2. Literature Review

In this section, the important features of AI Integrated Mobile Health applications will be presented, along with an evaluation on the integration of AI into the decision-making processes of organizational structures. The fundamental attributes of AI based Decision Support Systems will be outlined with regards to the literature examined. Research literature on the topics of AI Integrated Decision Support Systems for mobile health and wearable healthcare devices will be discussed.

2.1. Background about the AI-Powered DSS

Within the healthcare sector, decision support systems (DSS) are commonly implemented with the aim of refining professional judgment and patient care outcomes [7]. As indicated in **Figure 1**, the implementation of machine learning (ML) algorithms has enabled the DSSs to improve clinical outcomes. These systems aid in performing risk analysis, therapy determination, and even diagnosis through massive dataset computations using rule-based systems, algorithms, and predictive analytics [8]. This research discusses the application of AI algorithms in DSSs within the healthcare industry. AI integration in DSS refines suggestions accuracy by learning from historical and contemporaneous patient data. Automated analytical processes reduce healthcare professionals' cognitive workload while enhancing the accuracy of diagnoses. The continual evolution of AI-driven DSS enhances their predictive capabilities and strengthens medicine's evidence-based foundation.

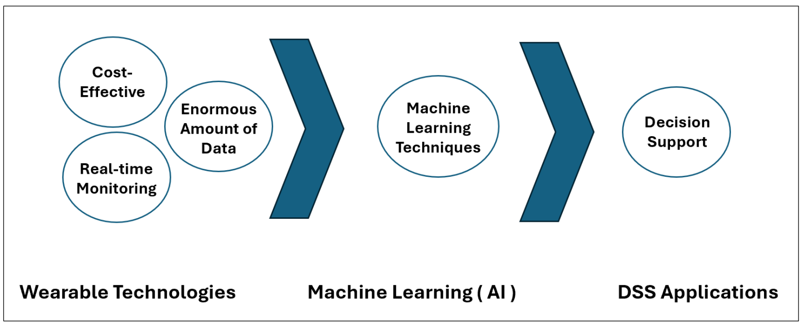

Figure 1. Sample of AI-Powered DSS applications.

The use of AI in healthcare has drastically changed patient monitoring, treatment planning, and illness diagnosis [9]. By detecting complicated patterns in medical data, machine learning algorithms improve diagnosis prediction and early illness identification. As AI technology develops, its use in healthcare is growing, motivating innovation in operational and clinical contexts.

Personalized healthcare and remote patient monitoring have been transformed by the emergence of wearable technology and m-Health. Continuous physiological data collection is made possible by mobile applications and sensor-equipped devices, which enable early intervention and real-time health tracking [10]. AI-powered analytics improve preventative care and chronic illness management by gaining valuable insights from wearable data. Proactive health monitoring is made possible by the growing use of wearable technology, which reduces the need for in-person healthcare appointments [11]. It is anticipated that as m-Health and AI integration continue to progress, healthcare accessibility and efficiency will be improved globally.

2.2. Characteristics of the AI-Powered DSS

AI-powered decision support systems (DSS) utilize extensive datasets alongside predictive analytics to optimize decision-making in medicine. These systems integrate machine learning algorithms to enhance therapy suggestions and increase the precision of diagnostics [12]. Automated reasoning capabilities provide far greater clinical evaluation accuracy by mitigating human error, which in turn enables the delivery of real-time insights. Context-aware computing incorporates ambient cues, vital signs, and patient history to better personalize medical recommendations. DSS keeps developing along with AI, providing medical personnel with ever-more-advanced tools for making accurate decisions.

m-Health applications serve as a bridge between patients and healthcare providers by enabling remote health monitoring and real-time data transmission, as presented in **Table 1**. These applications utilize AI algorithms to detect anomalies, track chronic conditions, and provide personalized health recommendations. Adaptive learning mechanisms refine user interactions, ensuring that health interventions are tailored to individual needs. Secure cloud-based infrastructures store and analyze patient data while maintaining privacy and compliance with regulatory frameworks. With ongoing technological advancements, m-Health applications are continuously improving patient engagement and accessibility to healthcare services [13].

Table 1. Capabilities of current AI models in m-Health and wearable devices applications.

Field	AI Integration	Real-Time	User Personalization	Data Secured
Remote Monitoring	✓	✓	✓	✓
AI-Based Diagnostics	✓	✓	✓	✓
Personalized Health Tracking	✓	✓	✓	✓
Chronic Disease Management	✓	✓	✓	✓
Anomaly Detection	✓	✓	✗	✓
Real-Time Alerts	✓	✓	✗	✓
User Engagement	✓	✓	✓	✗
Cloud-Based Storage	✓	✓	✗	✓
Cardiovascular Detection	✓	✓	✓	✓
Early Diagnosis & Prediction	✓	✓	✓	✓

Continuous physiological signal collection, analysis, and interpretation from sensor-equipped devices is known as wearable data processing. Early health risk identification is made possible by AI-driven algorithms that derive valuable insights from real-time physiological data. Through noise reduction, signal processing techniques improve data accuracy. By processing data locally on wearable devices, edge computing solutions reduce latency and reduce reliance on cloud-based computations. Wearable technology is predicted to further transform preventive medicine and personalized healthcare as it develops and integrates with AI.

2.3. Related Studies on AI-Integrated DSS, m-Health, and Wearable Healthcare Technologies

Several strategies were previously explored to manage and monitor different health disorders using multiple AI models, including machine learning. Wearable devices are widely used to gather physiological and behavioral data. Deep learning models are frequently applied to improve diagnostic accuracy. Figure 2 demonstrates the classification framework as a tree map to categorize the related studies about AI-Integrated DSS, m-Health, and Wearable Healthcare Technologies. This tree map classifies the recent studies reviewed in the literature according to their focus (m-Health, Wearable Technologies, AI-DSS) and the machine/deep learning models they used.

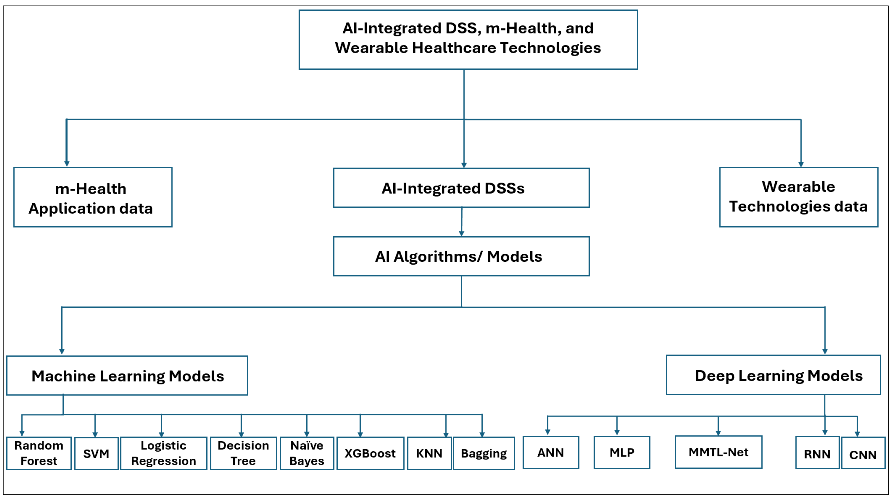

Figure 2. Tree map of recent studies on AI-Integrated DSS, m-Health, and wearable healthcare technologies.

Resistance monitoring of lower limb movements in real time is essential for several clinical and athletic applications, including athletic training and rehabilitation. The study by Burenbatu *et al.* introduced a unique Mobile Multi-Task Learning Network (MMTL-Net) architecture that combines MobileNetV3 for fast feature extraction and leverages multi-task learning to concurrently forecast resistance levels and detect activities. Experimental results reveal that MMTL-Net greatly outperforms current models using the UCI Human Activity Recognition and Wireless Sensor Data Mining Activity Prediction datasets, attaining a higher

Resistance Prediction Accuracy (RPA) of 91.2% [14].

The authors Abdulhussein & Bilgin conducted a comparative analysis of machine learning algorithms for predicting heart disease. They assessed the efficacy of various ML algorithms on 319,796 patient records utilizing the KNIME platform. Logistic regression attained the greatest level of accuracy (91.43%) among the models evaluated, highlighting the potential of machine learning in precise heart disease prediction to support clinical decision-making [15].

The study by Pal *et al.* proposed a comparison analysis of different machine learning classifiers such as Random Forest, Logistic Regression, Support Vector, Naïve Bayes, Decision Tree, and K-Nearest Neighbors. Experiments employed four datasets, all sourced from Kaggle. In the heart disease dataset, the best accuracy achieved was 82.35%. While the heart disease dataset 2020, the highest accuracy was 74.59% [16].

Health services customization and early risk prediction constitute the key research issues in m-health systems, which could be solved using the application of AI algorithms and tools applied to physiological and behavioral data acquired by wearables and IoT devices in real-world situations. The researchers Delmastro *et al.* suggested a summary of the findings from research activities that focused specifically on AI-powered m-health systems as support for individualized rehabilitation services and risk assessment for malnutrition, mobile sensing data analysis for disease detection, and the finding of novel behavioral and health markers that can facilitate clinical practice and remote patient monitoring. They achieved an accuracy with 94% and recall value with 92% results [17].

The integrated AI systems have enhanced health practitioner decision-making capabilities. AI-CDSS promotes in-CVD and healthcare sectors, with personalized therapy while integrating predictive analytics through deep learning models of CNNs and RNNs. The study by K & Abirami, has proposed an innovative AI-CDSS that offers advanced analytics and real-time monitoring functions, including diagnosing and treating cancer, managing chronic diseases, optimizing medication, supporting surgical decision making, managing infectious disease outbreaks, analyzing radiology and medical imaging, providing mental health support, as well as conducting clinical trials and research. Moreover, the study describes the methods that are currently applied, for instance, the use of deep learning frameworks like convolutional neural networks (CNNs) and recurrent neural networks (RNNs) for predicting the risk of cardiovascular disease. Nevertheless, it underscores the importance of reevaluation, thorough validation, and incorporation of ethical as well as regulatory aspects concerning the application of AI models in clinical practice [18].

Cardiovascular disease (CVD) is a chronic and deadly condition that causes the greatest number of deaths worldwide; it is the medical community's top concern. Wearable sensor devices have grown popular in the contemporary healthcare context. They have allowed real-time monitoring of health data, therefore assisting in the early detection of the risk of heart disease. The correct diagnosis and prognosis of cardiovascular illness are crucial in giving proper therapy to patients by cardiologists. The researchers Gola & Arya, proposed a model that may properly forecast cardiovascular disorders and therefore lower the mortality rates connected with them. The Satin Bowerbird optimization technique picks the most relevant feature, and an upgraded deep-learning model is applied for classification. Here the performance of the suggested work is compared with different algorithms such as SVM, Decision Tree, Logistic Regression, Random Forest, and their Evolutionary Deep Learning model. Metrics used are accuracy, recall, precision, and Fl-score were calculated to determine effectiveness. The proposed model attained an accuracy of 90% as best performance [19].

The review conducted by Bani Hani & Ahmad, focuses on examining and summarizing the most accurate machine learning classifier applied for forecasting ischemic heart disease (IHD). Utilizing numerous databases, an exhaustive investigation was implemented. Thirteen articles that were published between 2017 and 2021 were reviewed. The clinical outcomes to enhance the quality of care, the accuracy of algorithms to predict ischemic heart disease, and the commonly used algorithm to predict ischemic heart disease were the three themes obtained. The study finding shows that cardiovascular health disorders can be accurately predicted using machine learning algorithms such as DT, XGBoost, SVM, and LR [20].

The Authors S. Sharma & Singhal, developed a recommendation model based on manifestation that produced an impressive accuracy in predicting heart disease by using the Boosting ensemble approach. The dataset that is being used comes from a survey that was conducted in a group of 400,000 individuals. The purpose of the study was to collect basic information about health. The Personal Key Indicators of Heart Disease dataset has an uneven distribution of positive and negative classifications, which might hinder its effectiveness. Instances of the minority class were created by using oversampling pre-processing methods. The findings showed that Borderline 2 SMOTE-based XgBoost (Extreme Gradient Boosting) had the highest accuracy rate [21].

Machine learning has the potential to significantly contribute to the identification and prediction of potential risk factors for heart disease based on clinical and patient data, which is both dependable and cost-effective. Using survey data from 400,000 US residents, the researcher R. C. Das *et al.* developed and assessed six machine learning models for the prediction of heart disease. The six machine learning models that were evaluated and compared included Xgboost, Bagging, Random Forest, Decision Tree, K-Nearest Neighbor, and Naïve Bayes. The findings showed that Xgboost model were optimized, with an accuracy of 91.30% [22].

The human heart pumps oxygen-rich blood throughout the body with the assistance of coronary arteries. However, the probability of developing cardiac disease can be reduced through a proper diagnosis and early forecast. The study by Gangadhar *et al.* investigates the potential of predicting cardiac illness at an early stage using deep learning models. They explored early-stage prediction of coronary heart disease using ANN and other ML models like SVM, Random Forest, Decision Tree, and KNN. The findings showed that the ANN model demonstrated superior predictive accuracy of 88.44% [23].

Studies often focused on mood disorders and schizophrenia due to their overlapping complicated symptoms. The data collected from wearable devices were commonly employed in detection models which is producing promising results. The study by Nguyen *et al.* investigates the use of wearable device data used by multiple deep learning models to differentiate mental health disorders. The authors employ various deep learning algorithms, including CNN-based models and RNN-based models to analyze motion signal data from the Psykose (schizophrenia) and Depresjon (mood disorders) datasets. The performance evaluation show that the best-performing model was the XceptionNet model, achieving an accuracy of 86%. While for the task of differentiating between mood disorders and schizophrenia, the top-performing models were Resnet50v2 and LSTM, both achieving an accuracy of 88%. The study makes major contributions by demonstrating how using deep learning models with wearable technology can improve objectivity and accuracy in mental health diagnosis. The findings highlighted the possibility of such methods in clinical decision support [24].

Wearable devices have made it easier to gather data without having to enter it manually. The researchers Miah *et al.* conducted study that included making judgments for small or large sports teams on whether a given athlete was a suitable match for a specific game. They compared multiple machine learning algorithms to predict human behavior and health using data obtained from sensors put on patients. Five machine learning algorithms were utilized in their research to examine and predict human health behavior. XGBoost outperformed the other machine learning methods, achieving the highest accuracy [25].

The researchers Kuruvilla & Balaji, proposed a study utilizes a feature selection technique to determine the most suitable features of cardiac disease attributes. For their experiments, they used the Framingham heart disease dataset (FHS), which was collected from the Kaggle Machine Learning repository. Four ML classifiers validated 16 attributes in the dataset. Two feature selection methods, Correlation Feature Selection and Principal Component Analysis, were used to compare the study. The experimental result showed that the Correlation Feature Selection with Multilayer perceptron model obtained this dataset's highest accuracy of 84.9%. Demonstrated the effectiveness of Correlation Feature Selection in improving machine learning model performance for heart disease prediction [26].

Psychology Decision-making was initially dominated by early techniques. There are numerous efforts to bridge this gap by applying expert decision-making theories to fulfil users' needs for managing Parkinson's disease. The study by Timotijevic *et al.* developed a CDSS for Parkinson's disease using a theoretical framework suggested that Machine Learning could be used in the system to analyze patient data. The study was executed in four European countries to identify essential user needs and cognitive demands. The research methods used include Hierarchical Task Analysis (HTA), computational modelling, and vignette studies, drawing on interview and observational data collected from 47 prescribing clinicians across the four countries. The findings highlighted how the integration of data from various sources, such as self-reports data and data from devices. It can be able to enhance the decision-making process for clinicians to improve the management of Parkinson's disease. The study proposing a user-centered, theory-based framework for designing flexible, shared m-Health CDSS; by developing core principles for designing a Parkinson's m-Health based CDSS [27].

There have been different efforts made towards better insulin dosing in type 1 diabetes patients. For this specific case, decision support systems (DSS) were implemented with a great deal of machine learning expertise. The study by Tyler *et al.* analyzed the possibility of supporting the management of type 1 diabetes patients with an AI-driven DSS focusing on insulin prescription and overdose detection. Weekly insulin recommendations were made using a glucose-derived K-nearest neighbours algorithm (KNN) that was trained on both simulation and real-world data. The KNN-DSS achieved enhanced glycemic control, facilitated adjustment of insulin doses, and early error detection all benefiting the diabetes management. The study highlights the potential of AI in transforming diabetes care through data-driven decision-making [28].

Across the previous studies examined, an increasing trend towards the usage of wearable technologies and mobile health (m-Health) applications has emerged. The authors Burenbatu *et al.* and Nguyen *et al.* used deep learning include MMTL-Net, CNNs, and RNNs to recognise human activities and classify mental health disorders in real time using sensor-based data from wearable devices. Deep learning models, such as ANN and CNN by the researchers Gangadhar *et al.* and Miah *et al.*, were widely used to capture complicated non-linear patterns in physiological data, resulting in high accuracy rates in forecasting health risks. Furthermore, some researches Tyler *et al.* and Timotijevic *et al.* investigated the AI's ability to predict diseases and enhance clinical decision-making using recommendation systems, hence improving therapeutic results and patient care. Ethical concerns and the necessity for regulatory compliance were briefly mentioned by K & Abiram. It indicates a growing recognition of these difficulties in AI healthcare applications.

An analysis of the previous studies shows a consistent direction in applying artificial intelligence (AI), particularly machine learning (ML) models, to healthcare data for disease prediction and monitoring. A majority of the studies, including those by Abdulhussein & Bilgin, and Pal *et al.* have shown the effectiveness of traditional ML algorithms—such as Logistic Regression, Random Forest, K-Nearest Neighbours, and XGBoost—in successfully predicting cardiovascular diseases. Similarly, Gola & Ary and Bani Hani & Ahmad, highlighted the role of feature selection and data preprocessing in enhancing the accuracy of machine learning models for ischaemic heart disease prediction. Studies collectively emphasize the importance of data preprocessing techniques and robust machine learning models in increasing predictive healthcare technologies.

The studies differ in the range of data used, which ranged from real-world clinical datasets (e.g., Framingham Heart Study) to wearable sensor datasets (e.g., M-HEALTH) to survey datasets such as the CDC's Indicators of Heart Disease. Additionally, the depth of feature engineering and model interpretability also differed. The researchers Kuruvilla & Balaji, emphasised feature selection techniques to increase performance, whereas Gola & Arya, concentrated on optimising deep learning architectures without paying close attention to explainability. Furthermore, while K & Abirami, identified ethical and regulatory challenges, most research did not fully address these essential factors, limiting the real-world applicability of their findings.

While research has shown that AI can improve health monitoring and disease prediction, there are still numerous crucial gaps. First, most present research is disease-specific, focussing on singular health issues such as heart disease or mental health disorders rather than building a generalised, patient-centred model. This research addresses these gaps by implementing and comparing multiple robust ML algorithms to evaluate their predictive accuracy and efficiency. This research combines wearable sensor data from the m-Health dataset with survey-based indicators from the Indicators of Heart Disease dataset. To ensure a more effective AI-powered DSS, this research intends to enhance AI model interpretability, optimize predictive accuracy, and real-time adaptability by utilizing feature selection techniques.

It is observed that meaningful improvements have been made in health monitoring and wearable devices through the integration of AI. Learning models, such as deep neural networks and k-nearest neighbors, are utilized to interpret data obtained from these devices. However, traditional methods require additional enhancement in research. The M-HEALTH dataset includes various physiological and motion data, while the Indicators of Heart Disease dataset contains data from the Centers for Disease Control and Prevention (CDC) annual survey, covering the health information of about 400,000 adults. These datasets provide suitable context for exploring robust algorithms such as random forests, gradient boosting machines, and extreme gradient boosting (XGBoost). Therefore, this research aims to manage the gap in literature by applying advanced machine learning models to these two datasets, offering deeper understandings into physical activity and health monitoring.

3. Methodology

The research employs a machine-learning-based experimental study using real-world data. This is due to the ability of machine learning-based research to provide the benefits of efficient data handling, predictive analysis, and model evaluation using formal data structures, enabling the analysis of health tracking and risk forecasting with greater accuracy. An inductive methodology has been followed for the achievement of research objectives and the formation of solid conclusions.

This research is mainly based on quantitative data in which the experiment is using numerical signs measurements.

Figure 3 presents the research design flowchart of machine-learning-based models using real-world data. The research uses two datasets available online on Kaggle repository: M-HEALTH and Indicators of Heart Disease. These datasets contain measurements for physiological and behavioral data for health such as heart rate, level of activity, weight (BMI), smoking, and sleeping habits. The data will be pre-processed to remove noise, standardize variables, and extract relevant features for analysis. Key features for predicting health risks will be selected through feature selection methods, allowing the models to focus on clinically significant aspects that can be prioritized. To ensure fair evaluation of the models, both datasets are further divided into training and testing subsets. Random Forest, Gradient Boosting Machines (GBM), and XGBoost machine models are used in this research to classify and forecast heart risks. These models were chosen due to their established robustness and ability to handle high-dimensional and non-linear healthcare data. Random Forest reduces overfitting while maintaining interpretability, GBM sequentially improves predictive accuracy by correcting prior errors, and XGBoost offers computational efficiency and scalability for large datasets. Moreover, this research includes an experimental setting in which AI models are iteratively refined and assessed against multiple performance metrics, including accuracy, precision, recall, and F1-score.

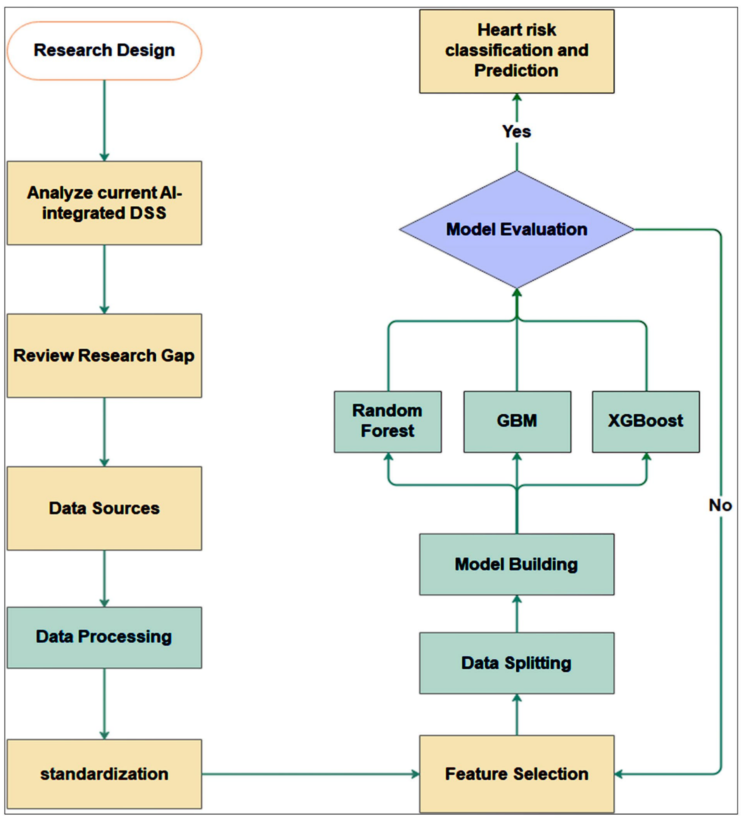

Figure 3. Research design flowchart.

3.1. Research Datasets

The target population of this research would be taken from two datasets M-HEALTH and Indicators of Heart Disease, available online in Kaggle repository.

The M-HEALTH dataset used in this research is based on records of body motion and physiological signals from 10 volunteers. The individual varied in profile and performed twelve different physical activities. The data is collected through three wearable Shimmer2 sensors worn on the chest, secured on the right wrist, and strapped on the left ankle using elastic bands to measure vital data of individuals. These sensors record motion data such as acceleration, gyroscopic movement, and magnetic field orientation, providing a complete view of the human body's movement as shown in Table 2.

Table 2. Dependent and independent variables attributes of M-HEALTH Dataset.

Dependent Variable	Independent Variables
Movement Activity	acc chest sensor (X, Y, Z) axes, gyro right lower arm sensor (X, Y, Z) axes, gyro left ankle sensor (X, Z) axes, acc left ankle sensor (X, Z) axes

The chest sensor also has the capability of bearing, though not included in the original model, always-ready ECG signals intended to be used in heart monitoring, with arrhythmia being a primary area of interest. All sensor readings are taken with a frequency of 50 Hz, which is appropriate for accurately observing variations in human activity. Activities have been captured varying from passive activities such as standing and sitting to actively running and jumping with the frequency and duration being specified. The files are separate for every subject, and recorded data are arranged in rows that reflect samples, while columns indicate readings or activities that were participated in over time. The activities were collected in an out-of-lab environment with no constraints on the way they must be executed, with the exception that the subjects should try their best when executing them. Thus, the data is collected in an uncontrolled manner, which is more applicable in real-life scenarios and increases variation proportionate to the manner of performance or execution.

Preprocessing encompasses cleaning, filtering, and standardizing sensor readings for use in machine learning, all of which are important for noise removal. This method enables the identification of familiar activities of daily living, facilitating the creation of health status monitoring and activity classification frameworks.

The Indicators of Heart Disease dataset, which can be found in the Kaggle repository, is useful in examining risk factors in different races of the American population [29]. Originally, the dataset, which aimed at heart disease risk factors encompassed around 300 variables, has been narrowed down to 18 key attributes for this study, having a target population of 319,795 patient records for this research. Features consist of dependent and independent variables, including body mass index (BMI), self-reported smoking, physical health, and diabetes presence, along with other variables, as shown in Table 3. These have a strong association with cardiovascular health (heart disease presence).

Table 3. Dependent and independent variables attributes of Indicators of Heart Disease Dataset.

Dependent Variables	Independent Variables
Heart disease presence	Age Category, DiffWalking, Stroke, Physical Health, Diabetic, Kidney Disease, Smoking, Physical Activity, Skin Cancer, Sex, BMI.

Preprocessing techniques will be applied to address missing information, redundant and duplicate data, and noise in the dataset, resulting in a suitable dataset for machine learning. Therefore, standardization and feature selection techniques will be used to improve the overall accuracy of the model, giving a simplified data structure that can get prediction. Imputation and cleanup of missing and noisy data will be performed to maintain the purpose of the dataset and conform to the specific data handling rules. The dataset serves as the basis for the study's predictive models. Additionally, the selected variables for analysis represent relevant risk factors for heart disease. It enhances the research value from both a clinical standpoint and its applicability across diverse populations.

3.2. Research Instrument

The online Google Colab was used as a research tool. Colab is a hosted Jupyter Notebook service that requires no setup and gives free access. It is highly suitable for machine learning and data science. Python will be the primary programming language due to its versatility and excellent data-handling libraries. In this research, the Pandas and NumPy libraries will be used for data analysis, which includes data transformation, cleansing, standardization, and pre-processing. The visual perception of the data patterns and distributions will be facilitated by the usage of more visualization libraries such as Matplotlib and Seaborn. Such visuals will also be useful during the first stage of data analysis when the features of the dataset will be explored to determine how different features will be related to each other.

3.3. Model Performance Evaluation

The model's performance will be assessed where the performance of the classification task will be measured which requires applying accuracy, precision, recall, and F1 score which provide different perspectives as discussed in the studies highlighted in the literature review. Accuracy measures the balance of correctly classified instances. Precision evaluates the reliability of positive classifications by calculating the ratio of correctly predicted positives to total predicted positives. At the same time, it makes it crucial when false positives are costly. Recall quantifies the model's ability to detect actual positive cases. It is essential in scenarios where missing positives have severe consequences. Finally, the F1-score is computed using precision and recall. It balances these two measures, providing a comprehensive performance evaluation.

Accuracy assesses the number of instances in a set of observations which are correctly classified. It can be computed in the next equation:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{Total instances}},$$

Precision refers to the measure of exact positive occurrences relative to the entirety of positive prediction. It can be computed in the next equation:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}},$$

Recall definition accesses the number of correctly predicted positive cases. It can be computed in the next equation:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}},$$

The F1-Score balances the asymmetry between Precision and Recall by calculating their harmonic-mean. It can be evaluated using the following equation:

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}.$$

where TP represents True Positives, TN are True Negatives. While FP represents False Positives rate and FN are False Negatives.

4. Results

This section focuses on analyzing results and findings of a detailed machine learning modeling analysis. After that, a comparison of the ML models' performance on the two datasets are presented.

4.1. m-Health Dataset Analysis

The m-Health dataset consists of 1,215,745 rows and 24 columns including the following features. Table 4 provides a summary of each feature type and description.

Table 4. The m-Health dataset features description.

Feature	Type	Description
Chest Acceleration (X, Y, Z)	Continuous	Acceleration data from the chest sensor along three axes.
Electrocardiogram (Lead 1, Lead 2)	Continuous	ECG signals recording heart activity.
Left-Ankle Acceleration (X, Y, Z)	Continuous	Acceleration data from the left-ankle sensor along three axes.
Left-Ankle Gyro (X, Y, Z)	Continuous	Gyroscope data from the left-ankle sensor measuring angular velocity.
Left-Ankle Magnetometer (X, Y, Z)	Continuous	Magnetic field data from the left-ankle sensor.
Right-Lower-Arm Acceleration (X, Y, Z)	Continuous	Acceleration data from the right-lower-arm sensor along three axes.

Continued

Right-Lower-Arm Gyro (X, Y, Z)	Continuous	Gyroscope data from the right-lower-arm sensor measuring angular velocity.
Right-Lower-Arm Magnetometer (X, Y, Z)	Continuous	Magnetic field data from the right-lower-arm sensor.
Label	Categorical	Activity classification (0 for null (not specified/completed activity), 12 activity labels).

The activity set includes the following activities:

- **L1:** Standing still (1 min)
- **L2:** Sitting and relaxing (1 min)
- **L3:** Lying down (1 min)
- **L4:** Walking (1 min)
- **L5:** Climbing stairs (1 min)
- **L6:** Waist bends forward (20x)
- **L7:** Frontal elevation of arms (20x)
- **L8:** Knees bending (crouching) (20x)
- **L9:** Cycling (1 min)
- **L10:** Jogging (1 min)
- **L11:** Running (1 min)
- **L12:** Jump front & back (20x)

NOTE: In brackets are the number of repetitions (Nx) or the duration of the exercises (min).

Correlation analysis was performed to identify relationships between numerical variables. A correlation matrix was generated and visualized as a heatmap using seaborn library in Python. This heatmap highlighted the strength and direction of linear relationships between variables, revealing potential associations to be further investigated.

The features correlation matrix for the m-Health dataset as shown in **Figure 4** indicates the degree of relationship between features of the datasets captured from the sensors. The correlation coefficients vary from −1 to 1, with dark red colors indicating strong positive correlations and dark blue colors indicating strong negative correlations. The highest correlations are found between signals captured from the same sensor type, indicating a high degree of dependency between axes contained in each device. In contrast, low correlations are observed across different sensor modalities, for instance, between accelerometers and magnetometers, suggesting relative independence between their data streams.

The lack of strength of most column correlation with the label more complicated due to the low correlation. elect_signal_lead1, elect_signal_lead2, mag_left_ankle_sensor_X_axis, mag_left_ankle_sensor_Y_axis, and mag_left_ankle_sensor_Z_axis demonstrate near-zero correlation with the response variable. Such features don't seem to add to a system's predictive performance, so they could be deleted for machine learning feature selection, which would improve computational efficiency without impacting the model's accuracy.

4.2. Indicators of Heart Disease Dataset Analysis

The Indicators of Heart Disease dataset consists of 319,795 rows and 18 columns. Table 5 provides a summary of each feature type and description.

The heatmap shows the correlation between the variables to highlight strongly related variables. The Seaborn Library in Python was used to visualize the heat map. The following Figure 5 shows the correlation between the dataset variables.

Figure 5 depicts a correlation heatmap that shows the correlations between several health indicators and heart disease. The colour gradient shows the degree and direction of correlations, with values closer to 1 or −1 suggesting stronger correlations. Age Category, Difficulty Walking, Stroke, Physical Health, and Diabetes have the strongest positive correlations with heart disease, indicating that they may be significant risk factors. Physical activity and general health have a negative correlation, suggesting a preventive impact. The heatmap also shows the interrelationships between variables, which may influence the predictive modelling and feature selection in heart disease study.

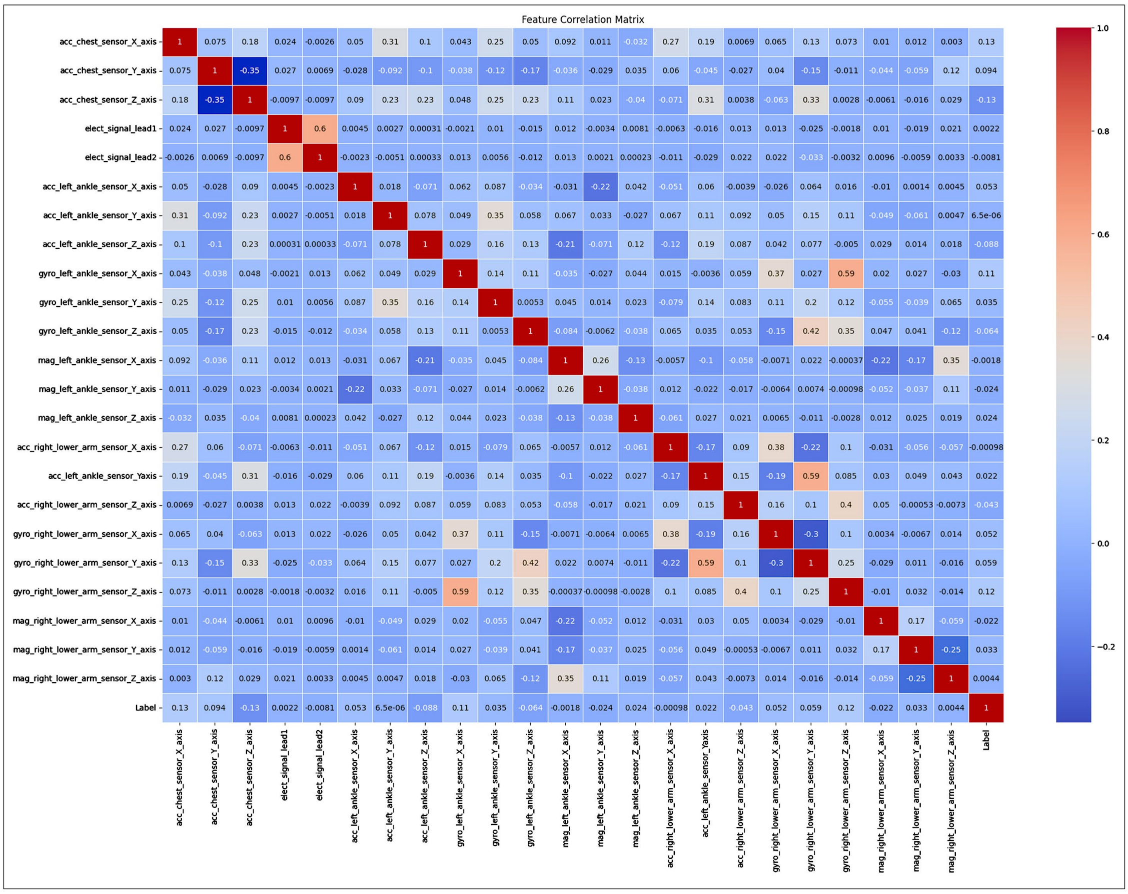

Figure 4. The m-Health dataset features correlation matrix.

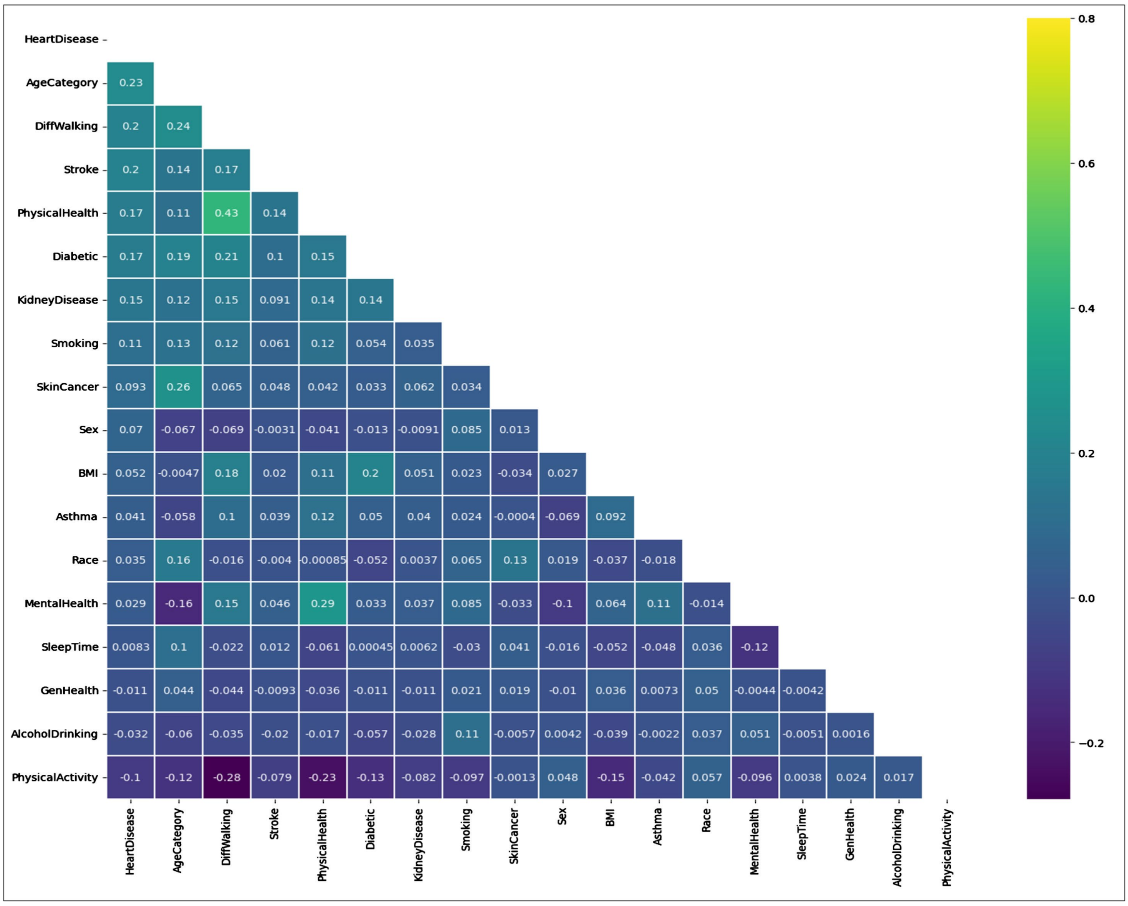

Figure 5. The indicators of heart disease dataset correlation heatmap.

Table 5. The heart disease dataset description.

Feature	Data Type	Description
Heart Disease	Binary	Presence of heart disease (Yes/No).
BMI	Continuous	Body Mass Index is a measure of body weight relative to height.
Smoking	Binary	Whether the respondent is a smoker (Yes/No).
Alcohol Drinking	Binary	Whether the respondent consumes alcohol (Yes/No).
Stroke	Binary	History of stroke occurrence (Yes/No).
Physical Health	Continuous	Number of days in a month with poor physical health.
Mental Health	Continuous	Number of days in a month with poor mental health.
Diff Walking	Binary	Difficulty in climbing stairs (Yes/No).
Sex	Binary	Gender of the respondent.
Age Category	Categorical	Age group classification.
Race	Categorical	Imputed race/ethnicity value.
Diabetes	Categorical	Whether the respondent has been diagnosed with diabetes.
Physical Activity	Binary	Engagement in physical activity outside of work.

Continued

Gen Health	Categorical	Self-reported general health (fair, good, very good).
Sleep Time	Continuous	Number of hours of sleep per day.
Asthma	Binary	History of asthma diagnosis (Yes/No).
Kidney Disease	Binary	History of kidney disease diagnosis (Yes/No).
Skin Cancer	Binary	History of skin cancer diagnosis (Yes/No).

4.3. Machine Learning Models

This section presents the findings of machine learning models applied to the m-Health and Indicators of Heart Disease datasets. The effectiveness of health monitoring and risk assessment was measured using each model's accuracy, precision, recall, and F1-score.

4.3.1. Machine Learning Models on m-Health Dataset

The m-Health dataset was analyzed with the Random Forest, GBM, and XGBoost models. The major goal is to determine the best model can be used to detect heart disease based on physiological data (senser-based), including Acceleration (chest and ankle) sensors, and Gyroscopic movement (arm and ankle) sensors.

Phase1: Feature Selection and Preprocessing

The correlation coefficient method was applied to identify the most critical features regarding a model's categorization accuracy and the most important factors impacting accuracy. Features with weak correlations ($|r| < 0.1$) were excluded. The selected variables were acc chest sensor (X, Y, Z) axes, gyro right lower arm sensor (X, Y, Z) axes, gyro left ankle sensor (X, Z) axes, acc left ankle sensor (X, Z) axes. Further, the dataset was split into two subsets, one for training which constituted 80% and the other one for testing which was 20%.

Phase 2: Models Performance Evaluation

The models were evaluated based on the weighted average of the four-performance metrics: accuracy, precision, recall, and F1-score.

As shown in Table 6, the Random Forest model achieved the highest score in all metrics of accuracy (94.23%), precision (94.36%), recall (94.23%) and F1-score (94.21%). This demonstrates that Random Forest maintains the equilibrium between precision and recall, yielding accurate predications with minimal false positives and false negatives. Its exceptional performance across all important criteria makes it the best model for correctly recognizing health-related activities in this dataset.

Table 6. The m-Health dataset models performance evaluation.

Model	Accuracy	Precision	Recall	F1-Score
Random Forest	**0.9423**	**0.9436**	**0.9423**	**0.9421**
Gradient Boosting	0.8365	0.8313	0.8365	0.8183
XGBoost	0.9034	0.9046	0.9034	0.9023

4.3.2. Machine Learning Models on Indicators of Heart Disease Dataset

The Indicators of Heart Disease dataset, which consists of health risk factors and behavioral data, was trained and tested using the same machine learning models. The main goal is to predict the presence of heart disease based on 11 highly correlated independent variables which include: Age Category, DiffWalking, Stroke, Physical Health, Diabetic, Kidney Disease, Smoking, Physical Activity, Skin Cancer, Sex, and BMI.

Phase1: Feature Selection and Preprocessing

Categorical variables were encoded correctly for model compatibility. In addition, Feature selection was conducted using a correlation coefficient approach. Variables with near-zero or weak correlations ($|r| < 0.1$) to the target were excluded, while highly correlated and clinically relevant predictors were retained. The selected variables were Age Category, Difficulty Walking, Stroke, Physical Health, Diabetes, Kidney Disease, Smoking, Physical Activity, Skin Cancer, Sex, and BMI. Moreover, the dataset was divided into 80% training and 20% testing subsets.

Phase 2: Models Performance Evaluation

The models were evaluated based on the weighted average of the four-performance metrics: accuracy, precision, recall, and F1-score.

As shown in Table 7, the gradient-boosting model outperformed all other models in terms of accuracy (91.26%), precision (92.03%), recall (91.26%), and F1-score (87.10%). These results show that Gradient Boosting efficiently balances precision and recall, making it the most accurate model.

Table 7. The ındicators of heart disease dataset models performance evaluation.

Model	Accuracy	Precision	Recall	F1-Score
Random Forest	0.8945	0.8505	0.8945	0.8689
Gradient Boosting	**0.9126**	**0.9203**	**0.9126**	**0.8710**
XGBoost	0.9123	0.8539	0.9123	0.8709

Furthermore, the recall is excellent, even slightly better than precision in some models, which suggests the models are able to detect positive cases effectively. While the F1-score for the heart disease dataset was slightly lower than that of the m-Health dataset, it remained in the high-performance category across all models. This minor discrepancy is consistent with the differences in feature complexity and data representation between the two datasets.

4.4. Integration of ML Models Performance (AI) with the DSS

In this section, the accuracy, precision, recall, and F1-score of Random Forest, GBM, and XGBoost on the m-Health and Indicators of Heart Disease datasets were compared to identify the most effective model for each dataset and assess their overall reliability and suitability for health-related predictions:

1) In terms of accuracy, Random Forest was best with the m-Health dataset

achieving 94.2%, whereas Gradient Boosting was best with the Indicators of Heart Disease dataset achieving 91.3%.

2) In terms of precision, Random Forest achieved the maximum precision of 94.4% with the M-HEALTH dataset, while Gradient Boosting was best with the Indicators of Heart Disease dataset achieving 92.0%.

3) In terms of recall, Random Forest was best with the m-Health dataset achieving 94.2%, while Gradient Boosting outperformed with the Indicators of Heart Disease dataset achieving 91.3%.

4) In terms of F1-score, Random Forest excelled in the m-Health dataset achieving 94.2%, while Gradient Boosting was best with the Indicators of Heart Disease dataset achieving 87.1%.

Overall, Random Forest is preferable for analysing activity data based on sensors, like the m-Health dataset. Gradient Boosting works better with more structured medical data, such as the Indicators of Heart Disease dataset. Ensemble-based models like Random Forest are quite effective for sensor data. As a result, aligning AI models with DSS frameworks will boost prediction accuracy. Thus, enabling more trustworthy and adaptive decision-making in digital healthcare.

5. Discussion

This section provides a thorough explanation and critical analysis of the results gained from the use of machine learning (ML) models in this research. The findings are connected to the research's hypothesis, objectives, and the broader context of AI integration in healthcare decision support systems (DSS). The results are further discussed in this section, looking at how the research's findings relate to previous research.

5.1. Discussion of the Hypothesis

The key hypothesis that guided this study was suggesting that integrating AI-driven (machine learning models) into clinical decision support systems might considerably increase the accuracy and reliability of health-related predictions. The null hypothesis argued that such integration would not result in any meaningful improvements/predictive accuracy.

In this research, the null hypothesis is examined deeply, questioning its validity considering the actual results of the experiment. Quantitative analyses were performed to evaluate whether the null hypothesis could endure the assumption that "applying a machine learning algorithm" at a dataset will not yield any meaningful predictions in terms of the underlying health issues involved.

In this case, the null hypothesis is examined closely with regard to whether the outcome from the experiments conducted would warrant an acceptance of its validity. The metrics of several performance evaluation measures were calculated to determine whether the null hypothesis is correct regarding whether the machine learning models cannot predict health issues with high accuracy. This research's null hypothesis stated that incorporating machine learning algorithms into health data analytics will not significantly increase the predictive performance for clinical decision-making.

This research's findings, based on extensive experimental study with Random Forest, Gradient Boosting, and XGBoost algorithms on two separate datasets which include the m-Health dataset (sensor-based) and the Indicators of Heart Disease dataset (survey-based) provide strong empirical evidence to **reject the null hypothesis**. Random Forest and Gradient Boosting exceeded baseline predictions, with accuracies exceeding 94% and 91% respectively.

Table 8 compares classification accuracies attained in recent studies on heart disease detection and human activity recognition to contextualise the suggested machine learning approach's performance. These studies use a variety of AI algorithms on a variety of datasets, displaying varying degrees of accuracy depending on methodology, dataset quality, and application breadth. The model in this research exceeds earlier studies, with a top accuracy of 94.23%, demonstrating its possibility for more reliable predictions in health monitoring DSS systems.

Table 8. Accuracy comparison of AI models in recent studies vs. this research.

Reference #	Recent Studies	AI Model	Model Accuracy
[14]	Burenbatu, Liu, Lyu (2025)	MMTL-Net Model	91.20%
[15]	Ayat, Bilgin (2024)	Logistic Regression Model	91,43%
[16]	Pal *et al.* (2024)	Decision Tree Model	82.35%
[22]	Das *et al.* (2023)	XGBoost Model	91.30%
	This Research	**Random Forest Model**	**94.23%**

This research is evaluated with respect to model evaluation including accuracy measure. In machine learning research, particularly in applied contexts like healthcare, hypothesis testing is frequently represented by performance indicators and model validation rather than traditional statistical significance testing. The metrics presented in the methodology and results chapters support the alternative hypothesis (H_1): when correctly trained and tested, AI (ML models) greatly enhance the health risk prediction accuracy.

The results also suggest that machine learning models can effectively handle both survey-based and sensor-based health data. This supports the hypothesis that such models can enhance the adaptability and prediction powers of decision support systems in a wide range of healthcare fields.

5.2. Comparison with Recent Literature

The performance of machine learning algorithms discussed in this research aligns with recent studies in the field of predictive healthcare analytics. The outcomes of this research are well correlated with several recent studies, Section 2.3. The study by Delmastro *et al.* endorses the employment of AI models for real-time healthcare monitoring and decision-making, particularly with data sourced from wearables and mobile devices.

Moreover, the impact of m-Health data on improving clinical outcomes has received a significant amount of focus. The researchers Delmastro *et al.* and Nguyen *et al.* pointed out the critical role of artificial intelligence in recognizing and analyzing behavioral and physiological data. This research makes use of the m-Health data set, which supports this hypothesis that sensor data, when processed with machine learning (ML) models, can indeed lead to automated real-time decision-making.

In other respects, the research is using ensemble machine learning (ML) models to wearable data, which integrates with, and in some cases broadens, the existing body of AI in Healthcare literature as discussed in recent works. This research uses of Random Forest, Gradient Boosting, and XGBoost resulted in top classification accuracy of 94%, which are competitive with and outperform similar work in the field.

This research uses ensemble models, which maintain interpretability while reaching competitive performance. Unlike deep neural networks, which can be computationally expensive and often opaque in clinical reasoning, the ensemble models in this research (Random Forest, Gradient Boosting) provide clearer feature importance outputs which is a critical component in clinician acceptance.

The researchers Pal *et al.* investigated different ML classifiers including Decision Trees, K-Nearest Neighbours, and Naïve Bayes across different datasets but obtained lower top accuracies of (82.35%) for heart disease prediction. This research's model performance (over 91%) demonstrates that ensemble approaches outperform basic classifiers such as KNN, Naïve Bayes, and Decision Trees. These findings emphasise the superiority of ensemble learners, particularly in structured health data, where this research models show greater performance without requiring sophisticated preprocessing or class rebalancing procedures.

Several studies have focused on hybrid optimisation strategies to improve model performance. The researchers Gola & Arya, developed a model that combined deep learning with the Satin Bowerbird algorithm. While this research does not use evolutionary algorithms, it does use robust ensemble learning to match these parameters, resulting in equalled or approached equivalent precision and recall without the need for sophisticated tuning, demonstrating their effectiveness and robustness in real-world clinical settings.

The study by S. Sharma & Singhal, used oversampling and SMOTE with XGBoost to handle the imbalance issue in clinical datasets in order to enhance prediction on imbalanced datasets. This research did not employ these methods. However, it still achieved high F1-scores, which suggests that ensemble models may effectively manage mild class imbalance due to mechanisms such as bootstrapped aggregation and regularisation. These models performed quite well, even in the absence of any synthetic data augmentation, which reflects moderate class imbalance.

The work of Burenbatu *et al.* introduced the Mobile Multi-Task Learning Network (MMTL-Net), which integrates MobileNetV3 feature extraction into multi-task learning for simultaneous activity recognition and resistance level estimation. Their experiments demonstrated that MMTL-Net achieved the highest performance on both the UCI Human Activity Recognition dataset and the Wireless Sensor Data Mining Activity Prediction dataset, with an accuracy of 91.2%. In contrast to MMTL-Net's complex, deep-learning multi-task architecture, this research focuses on simpler ensemble models that are more interpretable which is a crucial requirement in healthcare. Thus, demonstrating the robustness of ML models in providing consistent high accuracy across diverse datasets.

The research by Abdulhussein & Bilgin, performed a comparative study of the algorithms used in predicting heart diseases using machine learning. The researchers implemented ML techniques on a dataset of 319,796 patients using KNIME. Logistic regression showed the highest accuracy of 91.6%, which underscores the potential of machine learning algorithms for accurate predictions of heart diseases to aid clinical decisions. While their model is simpler than this research model, it lacks the power of ensemble learning provided by Gradient Boosting and the inability to model complex, nonlinear interactions which XGBoost handles elegantly; both of which derive this research's exceptional results.

The study by K & Abirami, reinforces the discussions by Tyler *et al.*, which underscore the role of AI in decision-making processes for chronic ailments and insulin administration. This research is consistent with existing studies while also providing a broader dual-dataset validation approach encompassing both real-time sensor data and population-level health indicators data which demonstrating greater applicability and generalisability.

Table 9 presents a comparison of existing AI approaches for health monitoring and heart disease detection, including their methodologies, results, and reported limitations.

Table 9. Comparison of this research and related recent studies.

Reference #	Authors	Year	Aims	Dataset Name	Methodology	Findings	Limitations
[14]	Burenbatu, Liu, Lyu	2025	Proposes AI-driven system for real-time lower limb resistance monitoring	UCI Human Activity Recognition dataset	MMTL-Net using MobileNetV3	Achieved 91.2% accuracy	It is not completed for heart disease prediction using wearable devices; challenges with real-time processing and sensor integration
[15]	Ayat, Bilgin	2024	Evaluates ML algorithms for heart disease prediction	Heart Disease Dataset (KNIME platform)	KNN, Naïve Bayes, Logistic Regression on 319,796 records	Logistic Regression: 91.6% accuracy	Limited generalizability due to dataset specificity; needs more patient data
[16]	Pal *et al.*	2024	Compares ML classifiers for heart disease prediction	Kaggle Heart Disease Datasets	Various ML classifiers with four datasets	Highest accuracy of 82.35% on the heart disease dataset; 74.59% on heart disease 2020 dataset	Performance varies across different datasets

Continued

[22]	Das *et al.*	2023	Compares ML models for heart disease detection	Clinical Heart Disease Dataset	Six ML models for clinical data	XGBoost: 91.3% accuracy	Requires more diverse data; overfitting on smaller data
	This Research	**2025**	**Proposes machine learning models using wearable and survey datasets**	**m-Health, Indicators of Heart Disease**	**Random Forest, Gradient Boosting, XGBoost**	**m-Health: Random Forest: 94.23%, Heart Disease: Gradient Boosting: 91.26%**	**Requires further optimization and generalizability using cross-validation methods to be improved**

As a result, this research provides compelling evidence to reject the null hypothesis, highlighting the fact that ML-based DSS integration will improve clinical predictions by combining constant high accuracy with wide applicability and interpretability. Furthermore, these research findings overcome the findings observed in the literature while also presenting a vital alternative: simpler, transparent models can perform like or better than, complicated, black-box systems in real-world clinical data contexts. Ensemble models provide feature importance scores, offering transparency and trust compared to "black-box" deep learning models, which is vital for clinician acceptance.

To summarize, this research supports the efficacy of ensemble ML models as interpretable, efficient, and highly accurate alternatives to more complicated AI systems, as well as providing a new perspective on multi-modal health data integration in AI-driven clinical decision support systems. Despite their predictive strength, deploying ensemble models in real-time decision support systems introduces practical challenges. Random Forest and GBM require significant computational resources for inference, which may limit their scalability on low-power mobile or bedside devices. In contrast, XGBoost is optimized for faster execution, making it more suitable for deployment in latency-sensitive contexts. Balancing model accuracy with computational efficiency remains an important trade-off for practical adoption. While this research benefits from testing across two datasets, the generalizability of the findings remains limited. Validation on more diverse demographic groups and across additional health conditions is required before clinical deployment.

6. Conclusions

This research evaluated the performance of Random Forest, Gradient Boosting Machines, and XGBoost on two datasets: m-Health and Indicators of Heart Disease. The findings confirmed the effectiveness of these models in real-time health risk prediction using wearable sensor data.

The findings of this research provide a significant contribution in terms of research and in terms of practice in the field of AI-integrated Decision Support Systems in the healthcare sector. Consequently, this research in terms of research was found to provide an empirical foundation for incorporating wearable data into clinical decision-making models. Moreover, the dual validation through independent datasets enhanced the study's scientific rigor, revealing how algorithm performance varies with different types of data including behavioral versus physiological. Furthermore, the feature selection improved model interpretability, aligning with the growing trend of explainable AI (XAI) in healthcare.

While in terms of practice, this research demonstrated the potential of sensor-derived data in clinical environments, emphasizing its feasibility for early diagnosis and real-time monitoring. Moreover, the model's performance outperformed previous studies, with a top accuracy of 94.23%, suggesting its capacity for reliable predictions in DSS. Furthermore, the results highlight that AI-powered DSS, when developed transparently and with diverse data, could significantly improve clinical outcomes. Future research should explore broader dataset integration, real-world testing, and continued advancements in explainability to further enhance AI-driven healthcare solutions.

6.1. Recommendations and Further Research

Based on the research's findings, several recommendations for future research are suggested to increase the usability and effect of AI-powered decision support systems in healthcare. Firstly, to assess the models' stability and generalisability, cross-validation techniques should be applied over several folds and data splits. Secondly, future study should look into class imbalance handling methods including Synthetic Minority Over-sampling Technique (SMOTE), under sampling, or weighted loss techniques. They help balance the learning process without affecting the performance, especially with under-represented specific diseases within medical datasets. Thirdly, future studies should strive to build upon this work by attempting to increase the breadth of the dataset to capture a wider spectrum of diseases, as well as the patient populations. This would enhance the model generalization and contribute toward constructing a more universal AI framework designed for diverse medical applications. Such models could allow for dynamic fine-tuning based on new patient information, thereby evolving the system's diagnostic capabilities over time. In addition, the use of multiple dataset types, including m-Health data and data from wearable sensors, could assist in developing more generalized models that enhance predictions made by AI-based DSS systems. Moreover, future studies should examine the clinically practical implementation of AI models within workflow processes, ensuring they are designed ethically, respectfully, and in a usable form for clinicians via explainable and interpretable methodologies. This entails comprehensive direct monitoring, active closed loop controlled cyclic feedback, and adherence to a medical standard. Likewise, the implementation of the models proposed in this research will necessitate the use of advanced techniques on data management including cloud-based secure storage, real-time analytics, and data integration from multiple secondary sources. These technical improvements will aid scalability, dependability, and incorporation into digital health systems. Furthermore, priority should be given to the design and implementation of a fully working, DSS interface. This could involve developing decision support system that provide predictions and suggestions directly to users. Integrating real-time sensor data from wearable devices will enable continuous patient monitoring, alarm generating, and context-sensitive intervention. As a result, this research lays the groundwork for a scalable, interpretable AI-DSS framework in healthcare. However, more effort is needed to validate these models in dynamic clinical real-world settings and expand their capability for broader use in real-time decision-making. In conclusion, while this research provides a solid foundation for integrating AI into decision-support systems using wearable data, it also suggests a variety of further investigations. The ultimate goal is to transfer from predictive modelling to intelligent, patient-centric, and automated decision support systems that are both clinically effective and technologically scalable.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Oresti, B. (2022) m-Health. https://www.kaggle.com/datasets/nirmalsankalana/m-Health-dataset-data-set-csv/data

[2] (2022) Indicators of Heart Disease. https://www.kaggle.com/datasets/kamilpytlak/personal-key-indicators-of-heart-disease

[3] Borges do Nascimento, I.J., Abdulazeem, H., Vasanthan, L.T., Martinez, E.Z., Zucoloto, M.L., Østengaard, L., *et al.* (2023) Barriers and Facilitators to Utilizing Digital Health Technologies by Healthcare Professionals. *npj Digital Medicine*, **6**, Article No. 16. https://doi.org/10.1038/s41746-023-00899-4

[4] Sáez, C., Ferri, P. and García-Gómez, J.M. (2024) Resilient Artificial Intelligence in Health: Synthesis and Research Agenda toward Next-Generation Trustworthy Clinical Decision Support. *Journal of Medical Internet Research*, **26**, e50295. https://doi.org/10.2196/50295

[5] Kazanskiy, N.L., Khonina, S.N. and Butt, M.A. (2024) A Review on Flexible Wearables—Recent Developments in Non-Invasive Continuous Health Monitoring. *Sensors and Actuators A*: *Physical*, **366**, Article ID: 114993. https://doi.org/10.1016/j.sna.2023.114993

[6] Lin, T., Chung, H., Jian, M., Chang, C., Perng, C., Liao, G., *et al.* (2024) An Advanced Machine Learning Model for a Web-Based Artificial Intelligence-Based Clinical Decision Support System Application: Model Development and Validation Study. *Journal of Medical Internet Research*, **26**, e56022. https://doi.org/10.2196/56022

[7] Chen, Z., Liang, N., Zhang, H., Li, H., Yang, Y., Zong, X., *et al.* (2023) Harnessing the Power of Clinical Decision Support Systems: Challenges and Opportunities. *Open Heart*, **10**, e002432. https://doi.org/10.1136/openhrt-2023-002432

[8] Khan, N., Nwafor Okoli, C., Ekpin, V., Attai, K., Chukwudi, N., Sabi, H., *et al.* (2023) Adoption and Utilization of Medical Decision Support Systems in the Diagnosis of Febrile Diseases: A Systematic Literature Review. *Expert Systems with Applications*,

220, Article ID: 119638. https://doi.org/10.1016/j.eswa.2023.119638

[9] Alowais, S.A., Alghamdi, S.S., Alsuhebany, N., Alqahtani, T., Alshaya, A.I., Almohareb, S.N., *et al.* (2023) Revolutionizing Healthcare: The Role of Artificial Intelligence in Clinical Practice. *BMC Medical Education*, **23**, Article No. 689. https://doi.org/10.1186/s12909-023-04698-z

[10] Restrepo-Parra, E., Ariza-Colpas, P.P., Torres-Bonilla, L.V., Piñeres-Melo, M.A., Urina-Triana, M.A. and Butt-Aziz, S. (2024) Home Monitoring Tools to Support Tracking Patients with Cardio-Cerebrovascular Diseases: Scientometric Review. *IoT*, **5**, 524-559. https://doi.org/10.3390/iot5030024

[11] Chaughule, S. and Bhonde, I. (2025) Influence of Remote Therapeutic Monitoring on Efficiency and Effectiveness of PT in Home Health Settings. *Medical Science*, **14**, p. 28. https://www.researchgate.net/profile/Shivangi-Chaughule-2/publication/388360184_Influence_of_Remote_Therapeutic_Monitoring_on_Efficiency_and_Effectiveness_of_PT_in_Home_Health_Settings/links/679414678311ce680c337e20/Influence-of-Re-mote-Therapeutic-Monitoring-on-Efficiency-and-Effectiveness-of-PT-in-Home-Health-Settings.pdf

[12] Iyortsuun, N.K., Kim, S., Jhon, M., Yang, H. and Pant, S. (2023) A Review of Machine Learning and Deep Learning Approaches on Mental Health Diagnosis. *Healthcare*, **11**, Article 285. https://doi.org/10.3390/healthcare11030285

[13] Kanakaprabha, S., Kumar, G.G., Reddy, B.P., Raju, Y.R. and Rai, P.C.M. (2024) Wearable Devices and Health Monitoring: Big Data and AI for Remote Patient Care. In: Sharma, N., Cengiz, K. and Chatterjee, P., Eds., *Intelligent Data Analytics for Bio-Informatics and Biomedical Systems*, Wiley, 291-311. https://doi.org/10.1002/9781394270910.ch12

[14] Burenbatu, Liu, Y. and Lyu, T. (2025) Real-Time Monitoring of Lower Limb Movement Resistance Based on Deep Learning. *Alexandria Engineering Journal*, **111**, 136-147. https://doi.org/10.1016/j.aej.2024.09.031

[15] Abdulhussein, A.B. and Bilgin, T.T. (2024) Comparison of Machine Learning Algorithms for Heart Disease Prediction. *İstanbul Ticaret Üniversitesi Teknoloji ve Uygulamalı Bilimler Dergisi*, **7**, 133-146. https://doi.org/10.56809/icujtas.1433853

[16] Pal, K., Panwar, S. and Choudhury, D. (2024) A Pragmatic Approach of Heart and Liver Disease Prediction Using Machine Learning Classifiers. 2024 *International Conference on Emerging Systems and Intelligent Computing* (*ESIC*), Bhubaneswar, 9-10 February 2024, 728-734. https://doi.org/10.1109/esic60604.2024.10481536

[17] Delmastro, F., Di Martino, F. and Campana, M.G. (2023) AI-Empowered m-Health Systems for Personalised Services and Digital Phenotyping. *Ital-IA* 2023: 3*rd National Conference on Artificial Intelligence*, *organized by CINI*, Pisa, 29-31 May 2023, 217-222. https://ceur-ws.org/Vol-3486/103.pdf

[18] Poomari Durga, K. and Abirami, M.S. (2023) AI Clinical Decision Support System (AI-CDSS) for Cardiovascular Diseases. 2023 *International Conference on Computer Science and Emerging Technologies* (*CSET*), Bangalore, 10-12 October 2023, 1-7. https://doi.org/10.1109/cset58993.2023.10346885

[19] Gola, K.K. and Arya, S. (2023) Satin Bowerbird Optimization-Based Classification Model for Heart Disease Prediction Using Deep Learning in E-healthcare. 2023 *IEEE*/*ACM* 23*rd International Symposium on Cluster*, *Cloud and Internet Computing Workshops* (*CCGridW*), Bangalore, 1-4 May 2023, 296-298. https://doi.org/10.1109/ccgridw59191.2023.00063

[20] Bani Hani, S.H. and Ahmad, M.M. (2023) Machine-Learning Algorithms for Is-

chemic Heart Disease Prediction: A Systematic Review. *Current Cardiology Reviews*, **19**, e090622205797. https://doi.org/10.2174/1573403x18666220609123053

[21] Sharma, S. and Singhal, A. (2024) A Model Based on Borderline 2 Synthetic Minority Oversampling Technique and Extreme Gradient Boosting for Predicting Cardiovascular Disease. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4994847

[22] Das, R.C., Das, M.C., Hossain, M.A., Rahman, M.A., Hossen, M.H. and Hasan, R. (2023) Heart Disease Detection Using Ml. 2023 *IEEE 13th Annual Computing and Communication Workshop and Conference* (*CCWC*), Las Vegas, 8-11 March 2023, 983-987. https://doi.org/10.1109/ccwc57344.2023.10099294

[23] Gangadhar, M.S., Sai, K.V.S., Kumar, S.H.S., Kumar, K.A., Kavitha, M. and Aravinth, S.S. (2023) Machine Learning and Deep Learning Techniques on Accurate Risk Prediction of Coronary Heart Disease. 2023 *7th International Conference on Computing Methodologies and Communication* (*ICCMC*), Erode, 23-25 February 2023, 227-232. https://doi.org/10.1109/iccmc56507.2023.10083756

[24] Nguyen, D., Chan, C., Li, A.A., Phan, D. and Lan, C. (2022) Decision Support System for the Differentiation of Schizophrenia and Mood Disorders Using Multiple Deep Learning Models on Wearable Devices Data. *Health Informatics Journal*, **28**, 1-23. https://doi.org/10.1177/14604582221137537

[25] Miah, J., Mamun, M., Rahman, M.M., Mahmud, M.I., Ahmad, S. and Nasir, M.H.B. (2022) MHfit: Mobile Health Data for Predicting Athletics Fitness Using Machine Learning Models. 2022 *2nd International Seminar on Machine Learning*, *Optimization*, *and Data Science* (*ISMODE*), Jakarta, 22-23 December 2022, 584-589. https://doi.org/10.1109/ismode56940.2022.10180967

[26] Kuruvilla, A.M. and Balaji, N.V. (2021) Heart Disease Prediction System Using Correlation Based Feature Selection with Multilayer Perceptron Approach. *IOP Conference Series*: *Materials Science and Engineering*, **1085**, Article ID: 012028. https://doi.org/10.1088/1757-899x/1085/1/012028

[27] Timotijevic, L., Hodgkins, C.E., Banks, A., Rusconi, P., Egan, B., Peacock, M., *et al.* (2020) Designing a m-Health Clinical Decision Support System for Parkinson's Disease: A Theoretically Grounded User Needs Approach. *BMC Medical Informatics and Decision Making*, **20**, Article No. 34. https://doi.org/10.1186/s12911-020-1027-1

[28] Tyler, N.S., Mosquera-Lopez, C.M., Wilson, L.M., Dodier, R.H., Branigan, D.L., Gabo, V.B., *et al.* (2020) An Artificial Intelligence Decision Support System for the Management of Type 1 Diabetes. *Nature Metabolism*, **2**, 612-619. https://doi.org/10.1038/s42255-020-0212-y

[29] Simmons II, B. (2021) Investigating Heart Disease Datasets and Building Predictive Models. Ph.D. Thesis, Elizabeth City State University. https://libres.uncg.edu/ir/ecsu/f/Brandon_Simmons_Thesis-Final.pdf