**University of Zawia**

**Faculty of sciences**

**Computer Science Department**

**Postgraduate Program**

# Systematic Review Personal Health Monitoring Algorithms Using Artificial Intelligence

*Presented By:* ***NADAY ALMUKHTAR ABOUBGHEELAH***

*Supervisor Name:* ***Dr. FATHI ABUOJAYLAH HAMHOUM***

***Academic Year***

**A thesis submitted to the Faculty of sciences University of Zawia in fulfillment of the requirements for the degree of Master of Science in the Department of Computer Science.**

27/07/2025

# الإقرار

أقرّ أنا ........................................، بأن ما اشتملت عليه الرسالة إنما هو نتاج جهدي الخاص، باستثناء ما تمت الإشارة إليه حتمًا وورد، وأن هذه الرسالة ككل أو أي جزءٍ منها لم يقدّم من قبل لنيل أي درجة علمية، أو بحث علمي لدى أي مؤسسة تعليمية أو بحثية أخرى، وللجامعة حق توظيف الرسالة أو الأطروحة والاستفادة منها مصدرًا مرجعيًا للمعلومات، لأغراض الاطلاع أو الإعارة أو النشر بما لا يتعارض وحقوق الملكية الفكرية المقررة بالتشريعات النافذة.

التوقيع.............................. :

التاريخ: ......./......./20م

بِسۡمِ ٱللَّهِ ٱلرَّحۡمَٰنِ ٱلرَّحِيمِ

﴿قَالَ رَبِّ اشْرَحْ لِي صَدْرِي (25) وَيَسِّرْ لِي أَمْرِي
(26) وَاحْلُلْ عُقْدَةً مِّن لِّسَانِي (27) يَفْقَهُوا قَوْلِي (28)﴾

صدق الله العظيم

سورة طه / الآية "25-29"

# Dedication

I dedicate this thesis to my dear parents, who have always been a source of inspiration and support. You are the reason for everything I am today, and your unconditional love remains my driving force to achieve the best. To my beloved siblings, who have always been by my side in every moment, your support and solidarity are priceless. To my friends and colleagues, who have been wonderful companions throughout this journey, thank you for your continuous encouragement and understanding. All of you are the reason for my strength, and I would not have been able to complete this journey

# Abstract

Personal Health Monitoring (PHM) systems, powered by artificial intelligence (AI), play a growing role in enabling preventive and personalized care. Despite numerous studies on AI integration in PHM, the existing literature remains fragmented and inconsistent. This thesis presents a systematic review of fifteen peer-reviewed studies published between 2022 and 2025, aiming to classify and evaluate AI algorithms used in PHM applications.

The analysis covers machine learning, deep learning, and hybrid models across different health conditions, data types, and performance metrics. The findings reveal an increasing reliance on deep learning architectures such as CNN and CNN-LSTM for real-time analysis, while traditional models like SVM remain useful in constrained environments.

The review also highlights critical limitations in current research, including limited clinical validation and lack of interpretability. Based on the findings, the study offers recommendations for selecting appropriate algorithms, ensuring ethical implementation, and guiding future research toward explainable, efficient, and adaptable AI models.

This work contributes a structured reference for researchers and system designers, supporting the development of robust, user-centered PHM systems driven by AI.

# ملخص

**تُعد أنظمة المراقبة الصحية الشخصية (PHM) المعززة بالذكاء الاصطناعي من الأدوات المتنامية في دعم الرعاية الوقائية والشخصية. ورغم تعدد الدراسات التي تناولت دمج خوارزميات الذكاء الاصطناعي في هذه الأنظمة، إلا أن الأدبيات المتوفرة تتسم بالتشتت والتباين. تهدف هذه الرسالة إلى إجراء مراجعة منهجية لخمس عشرة دراسة محكّمة نُشرت بين عامي 2022 و2025، بهدف تصنيف وتقييم الخوارزميات المستخدمة في تطبيقات المراقبة الصحية.**

**شملت التحليلات نماذج التعلم الآلي، والتعلم العميق، والهجينة، وفقًا للحالات الصحية المستهدفة، وأنواع البيانات، ومقاييس الأداء. أظهرت النتائج اعتمادًا متزايدًا على نماذج التعلم العميق مثل CNN وCNN-LSTM، في حين تظل النماذج التقليدية كـ SVM فعالة في البيئات ذات الموارد المحدودة.**

**كما كشفت الدراسة عن عدد من التحديات، أبرزها ضعف التحقق السريري وصعوبة تفسير نتائج النماذج. وقد خلصت إلى مجموعة من التوصيات الخاصة باختيار الخوارزميات، وتطبيقها بشكل أخلاقي، وتوجيه البحوث المستقبلية نحو نماذج قابلة للتفسير وفعالة وقابلة للتكيف.**

**تُقدم هذه الرسالة مرجعًا منظمًا يُسهم في دعم تطوير أنظمة مراقبة صحية فعّالة ومتمركزة حول المستخدم بالاعتماد على الذكاء الاصطناعي.**

# TABLE OF CONTENTS

# List of Tables

# List of Figures

# Acronyms

| Acronym | Full Term |
|---|---|
| AI | Artificial Intelligence |
| CNN | Convolutional Neural Network |
| DL | Deep Learning |
| DT | Decision Trees |
| ECG | Electrocardiogram |
| EHR | Electronic Health Record |
| GBM | Gradient Boosting Machine |
| GDPR | General Data Protection Regulation |
| HIPAA | Health Insurance Portability and Accountability Act |
| HL7 FHIR | Health Level Seven - Fast Healthcare Interoperability Resources |
| IMDRF | International Medical Device Regulators Forum |
| IoT | Internet of Things |
| KNN | K-Nearest Neighbors |
| LIME | Local Interpretable Model-agnostic Explanations |
| LSTM | Long Short-Term Memory |
| mHealth | Mobile Health |
| ML | Machine Learning |
| PHM | Personal Health Monitoring |
| PPG | Photoplethysmography |
| PRISMA | Preferred Reporting Items for Systematic Reviews and Meta-Analyses |
| RNN | Recurrent Neural Network |
| SHAP | SHapley Additive Explanations |
| SMOTE | Synthetic Minority Oversampling Technique |
| SVM | Support Vector Machine |
| XAI | Explainable Artificial Intelligence |

# Chapter One
# Introduction

## 1.1 Introduction

In the present times, the integration of artificial intelligence (AI) technologies into healthcare has radically changed the approach of people in managing and monitoring their personal health. Among the most influential innovations in this regard are Personal Health Monitoring (PHM) systems that combine monitoring wearable devices, mobile applications and analytical data repositories that monitor biological metrics and behavioral indicators outside of conventional health care environments. These systems are in line with the movement towards preventive & personalised healthcare, to empower individuals for early intervention.

Artificial intelligence plays an important role in helping PHM systems to turn raw sensor data into actionable insights. Machine learning (ML) and Deep learning (DL) algorithms make it easy to find out the anomalies in the health system and to classify the health problem or the prediction of the health risk which helps to improve the responsiveness of the health system and gives the accuracy to the system, As the volume and variety of the health-related data will keep increasing, the effectiveness of the PHM system depends more and more on the choice and performance of the AI algorithms used for the system.

Despite the increasing number of research papers on AI in PHM, the field is still fragmented. Studies differ greatly in their algorithmic methods, application domains, methods for evaluation, and in the quality of reporting. While some of them focus on algorithmic accuracy, some have a clinical relevance, some focus on interpretability and some have a focus on scalability. This diversity underlines the need for a structured synthesis of the literature in order to determine dominant trends in the literature as well as strengths in methodology and areas where further investigation is needed.

This paper aims to answer this call by conducting a systematic review of algorithms of AI used for personal health monitoring. It is intended to categorize and compare algorithms and critically evaluate them according to type, application, data characteristics and performance. The review gives a better understanding of the state of current practices, and aids to the informed decisioning for the development of AI-powered technologies for health monitoring (Esteva et al., 2019).

## 1.2 Problem Statement

The growing use of AI in personal health monitoring (PHM) has resulted in a significant growth of studies that applied various machine learning and deep learning algorithms to analyse health-related data. However, the literature is heterogeneous, and exhibits a very wide heterogeneity in terms of the type of algorithms applied, the medical conditions considered and/or the reported evaluation measures. This diversity means that it is hard to arrive at an understanding of the most common AI models in use, and their performance across conditions of health monitoring, as well as the methodological aspects of strengths and limitations for each. Furthermore, without such a structured and comparative synthesis, researchers and the developers of such systems cannot make informed decisions when selecting suitable algorithms for personal health monitoring systems. To address this lack of information, a thorough and systematic review is required to categorise and analyse existing studies, compare the performance of algorithms, and identify current trends and challenges in the use of AI in personal health monitoring.

## 1.3 Research Aims

1. To understand the present status of AI algorithms used in PHM systems, in terms of their proliferation throughout health solutions, we examine and map the usage of algorithms, identifying which domains of health related data are used in PHM systems, which data patterns are used and in which context are they deployed.
2. Classification of commonly used AI algorithms on the basis of types and areas of applications in PHM systems
3. Assess and compare published performance metrics of AI algorithms across different health monitoring use cases by attending to accuracy, interpretability, and ground-truthiness
4. Identify critical trends, technical challenges and ethical considerations related to the integration of AI into PHM systems for the future research and practical development

## 1.4 Significance of the Study

This research has both theoretical and practical implications in the area of artificial intelligence applications in healthcare.

This study has theoretical and practical implications in the field of artificial intelligence applications in healthcare. From an academic perspective, it is useful to make a contribution to building up the fragmented literature by providing a systematic and critical synthesis of AI algorithms applied in PHM. By categorizing algorithms, comparing performance of algorithms and assessing the methodological quality of published studies the review provides a sound basis for researchers that want to expand existing knowledge, or find gaps in the field. Practically, the study has some insights that can be useful for system developers, healthcare practitioners and policy makers involved in the design or the implementation of an AI-based PHM solutions.

By illustrating the most successful algorithmic methods and any emerging trends and shortcomings, the findings should be useful for informing improved decision-making in regard to the design of intelligent health monitoring systems that are accurate, interpretable and of clinical significance.

## 1.5 Research Questions

This study aims to answer the following research questions:

1. What are the most used artificial intelligence algorithms in personal health monitoring (PHM) systems?
2. What type, structure and application domain does these algorithms belong to?
3. What are the performance metrics that are usually reported when evaluating these algorithms?
4. How effective are the different AI algorithms for different PHM use cases?

## 1.6 Research Methodology

This study adopts a systematic review methodology to explore and synthesize the application of artificial intelligence (AI) algorithms in personal health monitoring (PHM) systems. The review process is structured according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, ensuring transparency, reproducibility, and methodological rigor.

The methodology includes the following key components:

- Formulation of research questions that guide the scope and analytical focus of the review.
- Definition of inclusion and exclusion criteria, which are designed to select relevant and high-quality studies. The inclusion criteria focus on peer-reviewed journal articles published between 2022 and 2025, written in English, and directly

addressing the use of AI algorithms (machine learning, deep learning, or hybrid models) for personal health monitoring tasks. Exclusion criteria eliminate non-peer-reviewed publications, preprints, non-English studies, and papers unrelated to personal health applications.

- Execution of a structured search strategy across multiple academic databases, including IEEE Xplore, PubMed, and ScienceDirect. Search strings include combinations of keywords such as "personal health monitoring," "artificial intelligence," "machine learning," and "wearable sensors."
- Implementation of a two-stage screening process, starting with a review of titles and abstracts to filter irrelevant studies, followed by full-text reviews to confirm methodological relevance and eligibility.
- Selection of studies based on the predefined criteria. For each selected article, relevant data are extracted regarding algorithm type, application domain, input data modality (e.g., ECG, PPG, accelerometer), performance metrics, and system implementation characteristics.
- Thematic synthesis and comparative analysis of the extracted data, which are organized into tables to highlight algorithmic trends, evaluation practices, deployment settings, and methodological patterns.
- Evaluation of ethical and practical aspects associated with AI integration in PHM, including explainability, user experience, and data privacy.
- Quality appraisal of the included studies using established evaluation tools to assess scientific robustness and reporting transparency.

## 1.7 Scope and limitations

This study focuses exclusively on the application of artificial intelligence (AI) algorithms in personal health monitoring (PHM) systems. Specifically, it examines peer-reviewed academic research, highlighting algorithmic approaches, performance metrics, deployment settings, and system-level implications.

The scope is limited to studies that apply AI—namely machine learning (ML), deep learning (DL), or hybrid models—to tasks such as anomaly detection, stress recognition, activity monitoring, or signal classification in PHM contexts. Only studies that utilize physiological or behavioral data from wearable or mobile health devices (e.g., ECG, PPG, accelerometers) are included.

The study excludes non-peer-reviewed literature, preprints, and non-English publications. It also omits studies focused purely on hardware design, clinical treatment evaluation, or hospital-based monitoring systems, as these fall outside the scope of AI-driven personal, remote health monitoring.

This delimitation ensures that the review remains tightly focused on algorithmic contributions and system-level implications relevant to the development and deployment of PHM applications powered by artificial intelligence.

## 1.8 Definition of Terms

To ensure conceptual clarity and consistency throughout this study, the following key terms are defined as used within the context of personal health monitoring and artificial intelligence:

1. Personal health monitoring (PHM): A class of digital health systems that continuously collect, analyze, and transmit physiological or behavioral data, typically via wearable or mobile sensors, to support self-care, early diagnosis, and management of chronic diseases.
2. Artificial intelligence (AI): A broad field of computer science that aims to emulate human intelligence in machines, enabling them to perform tasks such as learning, inference, classification, and prediction.
3. Machine learning (ML): A subset of AI that uses statistical techniques to enable computer systems to learn from data without the need for explicit programming. Examples include support vector machines (SVMs), random forests, and k-nearest neighbors (KNNs).
4. Deep learning (DL): A specialized class of machine learning techniques that uses multi-layer neural networks to learn from complex, large-scale data. Common architectures include convolutional neural networks (CNNs) and long short-term memory (LSTM) networks. • Hybrid models: AI architectures that combine two or more algorithmic approaches (such as integrating machine learning and deep learning techniques) to improve performance, robustness, and adaptability in specific applications.
5. Wearable devices: Electronic devices worn on the body (such as smartwatches and fitness bands) equipped with sensors that measure health indicators such as heart rate, movement, and oxygen saturation.
6. Electrocardiogram: A physiological signal that records the electrical activity of the heart over time, commonly used in PHM systems to monitor cardiac health.

7. Photoplethysmography: A non-invasive optical technique used to detect changes in blood volume in tissue microvasculature, often used in wearable heart rate and oxygen monitors.

These definitions provide a shared understanding of the terms used throughout the review and are aligned with their usage in the academic and clinical literature.

**Recall:** In the context of artificial intelligence and machine learning, recall measures a model's ability to correctly identify all relevant positive cases from the total actual positives in the dataset. It is calculated as the ratio of *true positives* to the sum of *true positives* and *false negatives*. High recall indicates that the model misses very few actual positive instances, which is especially critical in applications where failing to detect a positive case can lead to significant consequences, such as in medical diagnosis or safety monitoring systems.

## 1.9 Research Contribution

This thesis presents several original contributions to the area of AI applications in personal health monitoring (PHM). First, it provides a systematic and comprehensive synthesis of fifteen peer-reviewed studies which have not been collectively reviewed in the literature. By following a systematic review methodology and following the Protocol Reporting Items For Systematic Reviews and Meta-Analyses (PRISMA), the study fills a basic gap in the development of recent advancements in AI algorithms used in personal health monitoring systems. Second, to contribute to the literature by classifying and comparing various AI models such as machine learning, deep learning, and hybrid, regarding their performance metrics, feasibility of their deployment, and the type of input data in order to provide practical recommendations for AI model selection for developers, clinicians, and researchers aiming to decide which model should be chosen for real-world personal health monitoring applications. Third, the study sheds light on the technical and ethical issues surrounding the use of AI for personal health monitoring, including interpretability, dataset bias, and regulatory hurdles. Although these issues are often left out of technical papers, they are the key to ensuring that healthcare AI solutions are safe and equitable. Finally, the study provides new recommendations for future research directions, including the development of lightweight explainable AI models that can be used in resource-constrained environments and recommendations for improved interoperability and end-user experience. The work contains information on:

how to develop ethical, effective, user-centered personal health management (PHM) systems, making it a useful resource for both scholars and practitioners.

# Chapter Two
# Background and Literature Review

## 2.1 Overview of Personal Health Monitoring Systems

### 2.1.1 Definition and Scope of PHM Systems

Personal health monitoring (PHM) systems are a class of health technology which aim to continuously acquire, process and interpret physiological and behavioural information that fall outside of traditional clinical settings. Their aim is to help people monitor health status, identify signs of early disease and treat chronic conditions using digital technologies such as wearable sensors, mobile apps and smart devices (Pantelopoulos & Bourbakis, 2010). PHM is a paradigm shift in the delivery of healthcare from reactive and episodic towards proactive, continuous, preventive, self-managed and personalized healthcare. For the most part, PHM platforms are based on an integrated system architecture consisting of sensor technologies, data communication modules, data processing units and feedback systems. Sensors (heart-rate monitors, accelerometers, electrodermal activity detectors, etc.) collect data in its raw format, and they are then transmitted to a central system, a smartphone or a server in the cloud, for analysis.  This workflow provides for real-time monitoring and timely intervention which is very important in monitoring events such as diabetes, hypertension and cardiac arrhythmias (Buick et al., 2016).

The scope of personal health management systems (PHM) goes beyond medical management of specific diseases and includes health promotion and improvement of lifestyle. Consequently, these systems have started being used by healthy people for other reasons such as fitness, sleep analysis, and stress management among others, thus adopting the wider trend to healthcare consumerism. Moreover, lately, PHM platforms are integrated with telemedicine services, therefore facilitating remote consultations and creating the conditions for the longitudinal exchange of health data between the patient and the provider (Topol, 2015).

PHM systems are also driving broad based public health programs as they allow for the systematic collection of population wide data to inform epidemiologic studies and the development of public health policy. Nevertheless, the scope and efficacy of these systems is dependent on a number of determinants, such as device's accuracy, user's compliance, data privacy, and algorithm's performance - factors to be discussed in later sections of this review. Accordingly, the definition and boundaries of personal health management systems are inherently cross-disciplinary, involving biomedical engineering, data science, behavioral health and public policy. This complexity makes the need for a solid and scalable approach to system design and evaluation ever more evident, especially as artificial-intelligence technologies are becoming more widely integrated into PHM infrastructures.

### 2.1.2 Technological Components of PHM (Wearables, Sensors, Apps)

Personal health monitoring (PHM) systems are technology-based systems that consist of integrative hardware and software components designed to work in concert with each other for the continuous monitoring and systematic analysis of health-related data .

Typically, the technological architecture of personal health management systems consists of three main components: (a) wearable devices and sensors; (b) data sink and processing units; and (c) user-facing applications to support feedback provisioning and informative sources for clinical decision-making processes.

- **Wearable Devices and Sensors:**

Personal health management (PHM) systems are the focus of wearable technology proliferation at present, including wearable devices ranging from smartwatches, fitness bands, chest straps, to sensor-equipped apparel .

These devices involve a series of physiological sensors, such as photoplethysmographs (PPGs), electrocardiograms (ECGs), accelerometers (kinematic assessment), gyroscopes, thermistors, etc. which continuously collect data related to cardiac, locomotor, circadian sleep, body temperature, and other relevant physiological parameters (Patel et al., 2012). Their mobility and usability allow real-time continuous monitoring for extended periods of time in a wide variety of daily situations.

- **Data Transmission and Processing Units:**

Once collected, information gathered by the sensors is either sent directly to the intermediate platforms (via Bluetooth or Wi-FI) or transmitted to them (via Bluetooth or Wi-Fi), where they are either processed on the phone itself or on cloud servers. Most of these systems rely on AI to wash data clean, extract meaningful features, see deviations, identify trends, etc. Adding cloud computing helps to increase the scalability and allows us to process everything in real-time, even if the amount of data is huge (Swan, 2012). While we can eliminate some of that pre-processing on the device with edge locally (reducing latency and keeping data private), some of it will still need to be in the cloud.

- **Mobile Applications and User Interfaces:**

Mobile health (mHealth) applications serve as the user-facing layer of PHM systems. They provide visualizations of collected data, alerts for abnormal patterns, personalized recommendations, and interfaces for manual input of contextual data (e.g., mood, symptoms, diet). These apps often incorporate gamification elements to encourage user engagement and behavior change. Furthermore, many applications include secure channels for data sharing with healthcare providers, supporting remote diagnosis and chronic disease management (Marsch & Gustafson, 2013).

Interoperability between these components is a critical success factor for public health management (PHM) systems. Standards such as HL7 FHIR (Fast Healthcare Interoperability Resources) and IEEE 11073 facilitate data integration and communication between devices and health information systems. However, challenges remain related to standardization, energy efficiency, sensor accuracy, and cybersecurity, all of which require careful consideration during system design. The synergy of these technological components determines the operational capability of PHM systems and lays the foundation for integrating advanced AI models into subsequent stages of data analysis and decision support.

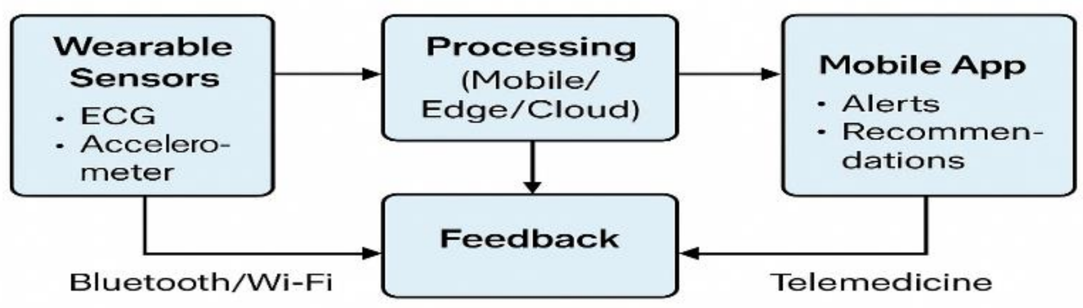


**Figure 1: Block Diagram of a Personal Health Monitoring (PHM) System**

**Source: Author (2025).**

The block diagram above illustrates the major components of a typical Personal Health Monitoring (PHM) system. It begins with those wearable sensors that collect all that body data stuff - like an ECG and accelerometers - these continue to pick up all that stuff. The raw signals then get sent over Bluetooth or Wi-Fi to some processing unit, which could be on your phone, on some local edge device, or in a cloud server. The layer of processing is responsible for cleaning the data, extracting useful data features and looking for anomalies using AI algorithms. After that, the processed data is transferred to a mobile health application, which can be considered as the front-end for you. The app offers real-time alerts and customized health tips and lets you manually enter data if you are so inclined. Furthermore, it also integrates with telemedicine, which allows information sharing with care providers or doctors in a secure manner. There's also a feedback loop, in that any user actions or input from the doctor go back into the processing system so it can fine-tune any future predictions and suggestions. That is, the closed-looped design helps the system to stay tuned in on each person's health patterns and remain useful over time.

### 2.1.3 Evolution and Milestones in PHM Development

The evolution of Personal Health Monitoring (PHM) systems has been tremendous during the past three decades, moving from the humble data-logger to the more advanced Artificial Intelligence (AI) driven cloud-enabled smart systems. Obviously there are some parallels with how technology and thinking in healthcare has transpired over the years, in particular, the shift to a more proactive and individualized approach to health-care - not unlike the discussions we have in class about where medicine of the future might take us.

- **Early Developments (1990s–2000s):**

relatively basic equipment for ambulatory monitoring, from Holter monitors to pedometers. Most of these tools were installed simply for offline data collection and subsequent crunching, so it wasn't real-time and a clinician was needed to pull the analysis. Fast forward to the early 2000s and you are starting to see mobile devices that In the 1990s, the first PHM projects emerged with are adopted to wireless networks which opened up a few remote monitoring possibilities - especially for patients suffering from heart or diabetes conditions (Korhonen et al., 2003).

- **Mobile Health and Wearables (2010s):**

The current decade was a big deal for PHM because smartphones and wearable devices like Fitbit, Apple Watch and other consumer health wearables let you access real-time data, sync with mobile computing platforms and perform simple health analytics right from your phone. With mHealth apps individuals could record their health behaviour, receive educational information and automated feedback

(Luxton et al., 2011). This period was also characterized by a great deal of interest in obtaining electronic health records. Electronic Health Records (EHRs) to get along well with user-generated data.

- **Next phase, AI and Cloud Implementation (late 2010s-present):**

Now, PHM is accelerating rapidly with the help of AI and cloud technology, Now, PHM is taking off at speeds aided by AI and cloud technology, and the use of machine and deep learning models is growing to help crunch the hard health data, find the abnormality and make a prediction (Rajkomar, Singh, Bender, Cerra, & Oh, 2019). Cloud-based systems allow one to store and access vast amounts of data, is multi-platform and can work through a great deal of data in a real time. That makes it easier to adapt to sophisticated PHM solutions inside or outside of hospitals.

- **Legislative Progress in Regulatory and Interoperability Standards:**

Apart from technological victories, there has been steps to the right direction in regards to regulation and standardization. PHM, however, has become safer and more ethical with the advent of the HL7 FHIR standard, FDA's approvals of some wearable medical devices and laws such as GDPR and HIPAA (Muoio, 2020).

As a result of these advances, PHM systems have become an integral part of the modern health care facility. They have the potential to help identify matters and problems early, to help and guide the longer-term care of disease, and to empower people to take responsibility for their own health. Still, the field is still evolving and we continue to encounter issues of data governance, clinical validation and data system integration that are addressed in the subsequent sections of this review.

### 2.1.4 Applications in Preventive and Personalized Healthcare

The fundamental shift that Personal Health Monitoring (PHM) systems are bringing to the structure of care delivery will be from a reactive approach to health issues to prevention and super-personalization of care. By being able to take in data all the time, and make use of AI enhanced analytics, these tools will allow students like me, and healthcare professionals to pick up when there is a potential serious health issue as it goes by

before they even present clinically, to tailor interventions for each individual and to improve long-term outcomes.

- **Applications of preventive Healthcare:**

One of the biggest successes of PHM is early identification of physiological, potential issues and risk factors. For example, keeping track of things such as heart rate variability, blood glucose or sleep patterns constantly give early warning signs for things like heart diseases, diabetes or sleep disorders. AI is meant to be a huge upgrade in this regard because unlike humans it can find subtle patterns that are easy to miss (Chung et al. 2019). In addition, PHM can send real-time alerts Or proactive suggestions for action (such as increasing your workouts, visiting a doctor, etc.) before things get out of hand.

- **Personalized Healthcare and Behavior Change**

Health machine personalization (PHM) also personalizes health suggestions and advice, based on the personal profile, habits, and context of the visitor or user. Advancing systems are powered by AI engines that intelligently tweak the goals, offer optimal activity levels and personalize the prompts on the basis of how you actually react and respond whilst measuring your biometrics. The individual feedback loop has been shown to keep us engaged with the mobile app and improve adherence (Banaee et al., 2013). In the treatment of chronic diseases, for example, these platforms enable treatments to be adjusted over time, making treatments more effective while minimizing the need to take unnecessary steps.

- **Population Level Benefits and Integration into Public Health Care:**

On a broader system level, PHM data that is aggregated can assist in guiding public health policy, identifying trends in risk factors enabling the initiation of targeted prevention interventions, and informing epidemiological models. For instance, during epidemics such as COVID-19, health apps and smart wearables could be used to observe symptoms, detect clusters, and observe recovery over time (Mishra et al., 2020). This shows that the reach of PHM is not only to optimise the health of individual people - it also contributes to the security of the entire health system. As cool as all this sounds, it is only the user retention in the heart of their services that determines the success of PHM in providing preventive and personalized care, data accuracy and finally, it is crucial to ensure model robustness clinically for AI models. The algorithms for personalization are continuously updated and controlled to be on the mark and valid, among different groups and conditions.

## 2.2 The Role of Artificial Intelligence in Prognostics and Health Management (PHM)

### 2.2.1 Core Functions of AI in PHM

Artificial Intelligence (AI) has become an integral component of Personal Health Monitoring (PHM) systems today and has enabled them to transcend the exquisite data collection to the active and intelligent interpretation and decision-making component. The three primary roles that AI plays in PHM, namely, detection, prediction, and classification, are essential to unlocking the value of continuous health data and the ability of the system to deliver timely, personalized, and clinically relevant insights.

**1 .Detection of Anomalies & Early Warning Indicators:**

Anomaly detection is one of the most important AI applications for PHM. By providing constant monitoring of physiological signals, like heart rate, respiratory rate, oxygen saturation and glucose levels, artificial intelligence (AI) algorithms can detect deviations from an individual's baseline or from what is considered to be normal for the population. Autoencoders or clustering techniques, both of which are unsupervised learning models, are popular choices to identify the outliers in time-series for the prediction of the occurrence of acute medical events such as arrhythmia, seizures, or hypoglycemia (Zhou et al., 2019) .

Different from static threshold based alerts that have been used in the traditional monitoring systems, AI based detection takes advantages of the contextual awareness. These systems have the potential for learning over time the patterns and correlations between multiple parameters that will enable them to detect complex physiological changes, which will function as predictors for a clinical event. For instance, subtle changes in sleep efficiency and increases in resting heart rate are abnormalities that appear to indicate the development of illness or a high level of stress .

**2 .Risk Stratification and prediction of Health Outcome**

AI models are constructed under predictive analytics for PHM to predict the health state in the future using historical and real-time data. This functionality is particularly valuable in chronic disease management, where these features can for example be of great use in predicting an exacerbation or a complication well in advance, which can create a significant difference in the patient's outcome and also lower the cost for the healthcare system. Machine learning techniques (random forest, gradient boosting machine, and recurrent neural network (RNN)) have been shown to be good at predicting hospital readmissions, disease progression, and even mental health deterioration (Rajkomar et al., 2018) .

AI - assisted risk stratification models are also used to categorize people based on their individual risk to develop certain conditions like hypertension or obesity or depression based on lifestyle, genetic predisposition and physiological trends. This allows intervention strategies to be individually tailored to the risk level of each individual, and therefore make preventative care more efficient.

**3.Classification of Health Outcomes and Diagnosis:**

Classification is probably the most used AI function in PHM, especially in the diagnostic and monitoring tasks. Supervised learning techniques such as support vector machine (SVM), k-nearest neighbor (KNN) and deep learning models (e.g., convolutional neural network or CNNs) are commonly used to classify sensor data into various health classes that are predefined. These can be either binary (e.g. normal vs. abnormal heart rhythm) or multi-class problems (e.g. sleep stage classification, activity recognition, or emotional state detection) (Dai et al., 2021).

A labeled dataset, which may be created through a clinical annotation or self-report, is often useful in a classification task. Once trained, the models can be used in real time, constantly processing the streams of data in order to apply health labels with high accuracy. This function is particularly important in mobile ECG monitoring, fall detection, and sleep quality monitoring.

- **Cross-Function Synergy and Adaptive Learning.**

For example, detection, classification and prediction are not separate functions but are mutually dependent. A PHM system may first detect an anomaly, which may be classified as a high-risk event, and then predict the probability of deterioration, initiating an intervention protocol. Furthermore, the use of AI systems gradually favors learning mechanisms that choose learning parameters There is often an adaptive mechanism which makes the machine learn parameters of the same model based on the occurrences of new data, user feedback or environmental changes for better performance and personalization in the long run.

The quality, granularity, and contextual richness of input data, however, have a significant relationship with the quality of your core AI functions; likewise, model interpretability and transparency have an impact on the quality of the results. Though

Although performance measures in terms of accuracy, precision, recall, area under the curve (AUC) have been published frequently, the explainability and clinical integration remain as issues of trust and adoption.

### 2.2.2 Integration of AI and IoT and Real Time Data Streams

The merger of Artificial Intelligence (AI) and Internet of Things (IoT) have been a significant force in driving the development of Personal Health Monitoring (PHM) systems. This integration enables the real-time collection of health data, transmission and analysis of that data along with feedback through an integrated network of connected devices and smart algorithms. It improves the responsiveness, scalability and personalization of PHM systems that are now transformed to dynamic tools for continuous health support.

- **IoT Infrastructure in PHM Systems**

The IoT ecosystem for health monitoring is comprised of a vast network of interconnected devices - incorporating but not restricted to wearable sensors, smart medical equipment, smartphones and home-based monitoring equipment - that produce streams of constant physiological and environmental data. These devices form a cyber-physical system that can be used to monitor various health parameters such as heart rate, oxygen saturation, temperature, physical activity, and even biochemical markers (Islam et al., 2015). They send information using wireless communication protocols like Bluetooth, Wi-fi or cellular networks to local gateways or cloud servers for processing.

- **Role of AI in Real Time Analysis:**

AI is the analytical engine that processes the raw data collected through the IoT devices. Through real-time data streaming platforms (e.g., Apache Kafka, Spark Streaming), machine learning models are used on data streams as they are created in order to provide extremely immediate interpretation and feedback. This functionality is needed for applications like fall detection in elderly care, epileptic seizure monitoring or arrhythmic episode detection in cardiovascular patients (Alam et al., 2018).

Moreover, machine learning can be applied to control over patterns and anomalies in time-series by using patterns like temporal convolutional networks (TCNs) or recurrent neural networks (RNNs). These models allow the PHM system to not only react to current conditions, but also to anticipate adverse events before they occur, in order to allow preventive intervention in real time..

- **Edge Computing for Low Latency Response:**

To overcome challenges related to cloud dependency, such as latency, data privacy and network availability, AI is increasingly being deployed at the edge, i.e. on the device or close to the data source. Edge-AI architectures enable the preliminary processing on smartwatches, mobile phones, or edge gateways that reduce the time for data interpretation and enhance the reliability of important health alerts (Shi et al., 2016). This is especially helpful in rural areas or areas with low connectivity.

- **Interoperability between Systems & Feedback Loops:**

Additionally, interoperability, such as device-to-device, data format, or communication protocols, can determine the success of AI-IoT integration. Protocols like MQTT (Message Queuing Telemetry Transport) and HL7 FHIR allow data to be shared between disparate systems and for decision-making to be coordinated and feedback loops provided to the end-user. Interactive analytics then use this information to provide contextual information - tuning thresholds for alerts, action suggestions, visualizing trends, and UIs, etc.

- **Security and Privacy Issues:**

The real-time integration of AI and IoT is a critical issue in terms of data privacy, security, and integrity. The AI models must be created in compliance with the healthcare data protection laws like HIPAA and GDPR. Wise to avoid centralizing sensitive information of users in training AI, new AI training methods are being explored, including federated learning and differential privacy (Yang et al., 2019) to ensure that users remain in compliance and trusted. To conclude, AI and IoT combined with real-time data streams allow PHM systems to possess continuous intelligence. This intersect supports an interesting, environmentally conscious and people-centric health ecosystem beyond the periodicity care model to live health management.

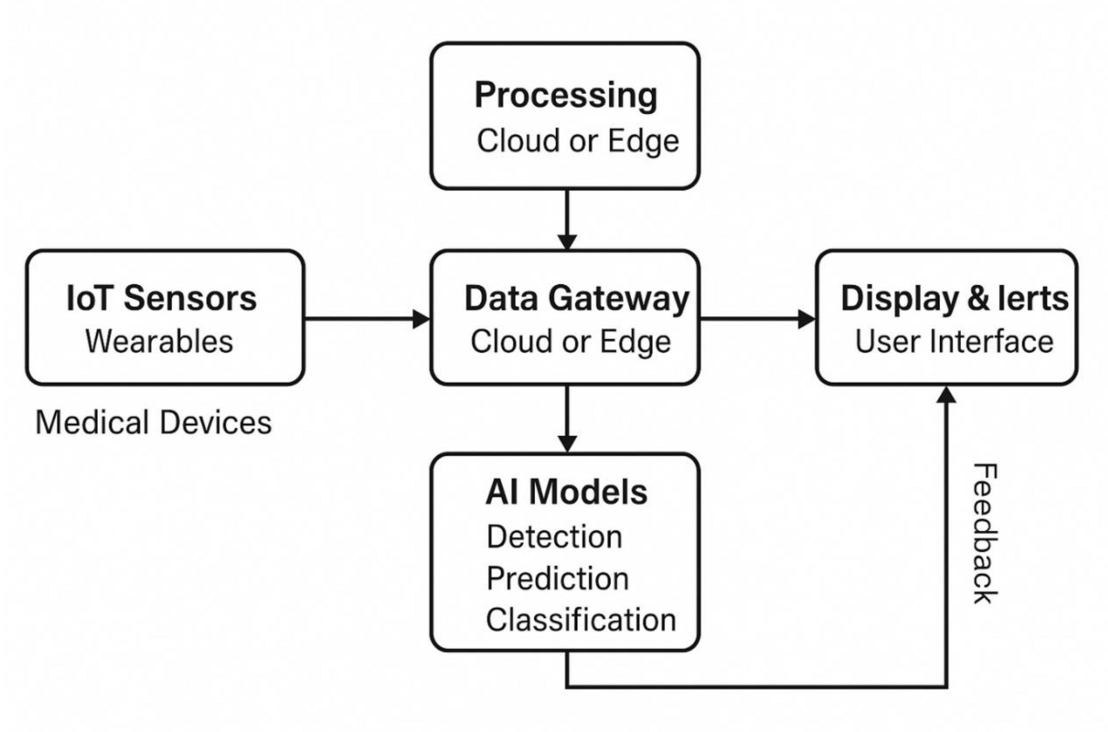


**Figure 2: Integration of AI with IoT and Real-Time Data Streams in PHM Systems**

**Source: Author (2025).**

This block diagram shows the integration of Artificial Intelligence (AI) and the Internet of Things (IoT) infrastructure in Personal Health Monitoring (PHM) systems. This entire process is started with IoT Sensors, which include wearable and medical devices, which acquire millions of physiological data every second. These data are sent to a Data Gateway, operating either in the cloud or at the edge, for first line processing. The gateway is interfaced with a Processing Unit, which acts to process and transmit data. Once processed, the data is passed to AI Models that are responsible for the detection, prediction, and classification of health events. The results are then passed to the Display & Alerts interface which provides real-time feedback and suggestions to the user. A feedback loop enables data gathered in a specific system to be fed back to the system for continuous learning and adaptive optimization of the sensor behavior and the AI outputs based upon user input or environmental changes..

### 2.2.3 Benefits of AI in Enhancing PHM Systems

The integration of Artificial Intelligence (AI) in Personal Health Monitoring (PHM) systems has significantly enhanced their functionality, reliability and clinical value. By making it possible to interpret data on the fly, provide individual feedback to the learner and even dynamically adapt the learning content, AI extends almost every aspect of PHM - changing it from passive data tracking to an active health management paradigm. The benefits of AI in PHM systems are multifaceted, encompassing technical, clinical, and user-centered dimensions.

**1. Improved Diagnostic Accuracy and Timeliness**

AI algorithms can analyze complex data with high dimensions related to health and detect anomalies and trends with high sensitivity and specificity. Machine learning (ML) and deep learning (DL) models in particular can find subtle changes in physiology patterns from normal values - often before they are visible clinically - therefore allowing for earlier intervention and more accurate diagnosis (Esteva et al., 2019). This level of diagnostic support can be particularly useful in patients who suffer from chronic diseases that can have their course of the disease and treatment outcomes greatly affected by early detection (e.g., diabetes, hypertension).

**2. Continuous, Real-Time Health Monitoring**

One of the most game-changing benefits of AI in PHM is that this technology can run continuously and in real-time. Unlike in traditional models of healthcare that are built around episodic visits to a doctor, AI-enhanced PHM systems offer around-the-clock surveillance of health parameters. Through data processing in real time and the identification of anomalies, these systems are able to notify those using the device, or healthcare facilities, of sudden problems such as cardiac arrhythmias, respiratory disorders, or sudden drops in oxygen level - and thus promote faster medical response (Litjens et al, 2017).

**3. Personalization of Health Insights and Interventions**

AI allows PHM systems to learn from an individual's health data over time and modify their models and suggestions according to each user's behavior, preferences and response. This dynamic personalization promotes the pertinence and the efficacy of health interventions. For example, AI can tailor activity objectives according to a user's baseline physical capability, optimise optimal medication times considering physiological variability or make dietary recommendations based on real-time trends of glucose (Ravi et al., 2017). Personalized feedback will help in enhancing the user engagement and adherence to health plans.

**4. Reduction of Healthcare Burden and Costs**

By bringing care home and allowing proactive disease management, AI-powered PHM systems can increase the frequency with which care is centered in people's homes and decrease emergency visits and hospital admissions as well as expensive late-stage interventions. Predictive algorithms make it possible to identify risks early and provide preventive care, which is much less costly than reactive treatment. Furthermore, AI can help healthcare providers to filter and prioritise patient data to improve clinical efficiency and lessen cognitive load (Topol, 2019).

**5. Support for Remote and Underserved Populations**

AI-based PHM systems can help address disparities in care. Generally speaking in rural or underserved communities. Through intelligent diagnostics, mobile These systems, including cloud-based platforms, interfaces, and mobile apps, can provide high-quality health. Monitoring in regions where there are no specialized healthcare structures AI also facilitates provide multilingual interface, automated triage mechanisms and decision support applications be used by community health workers with little training

**6. Data-Driven Research and Public Health Surveillance**

AI's ability to extract insights from large-scale PHM data opens opportunities for population-level research and health surveillance. Aggregated and anonymized datasets can be used to track epidemiological trends, assess the effectiveness of interventions, and inform public health policy. During the COVID-19 pandemic, for example, wearable data analyzed by AI helped predict outbreaks and monitor recovery trajectories (Mishra et al., 2020.(

Despite these benefits, the successful deployment of AI in PHM systems requires ongoing attention to issues such as algorithm transparency, data privacy, and regulatory compliance. Nonetheless, the potential of AI to enhance health outcomes, empower individuals, and transform healthcare delivery remains unprecedented..

## 2.3 Common AI Algorithms Used in PHM

### 2.3.1 Machine Learning Algorithms (e.g. SVM, KNN, Decision Trees)

Machine Learning (ML) algorithms are one of the most popular artificial Intelligence (AI) techniques for Personal Health Monitoring (PHM) systems. Their What gives them their strength is their ability to learn from data patterns and make data-driven decisions. making predictions or classifications without having to program them. These algorithms play a basics of processing health-related data produced by the wearable sensors and other monitoring devices. This part gives a brief of important ML algorithms. used in PHM, such as Support Vector Machines

(SVM), K-Nearest Neighbour KNN, Decision Trees (DT), with a focus on their applications and benefits. and limitations.

- **Support Vector Machines (SVM):**

Support Vector Machine (SVM) is a type of supervised learning algorithm, in which a hyperplane or set of hyperplanes are made in a high-dimensional space where the data points are classified. In the context of PHM, SVM is widely used for binary classification problems that can be related to medical sensors, like normal vs abnormal heart rhythms, normal vs stress level, identifying patterns of activity from sensor data and so on. Its capacity to work with high-dimensional datasets and to work well with limited samples makes it ideal for many health monitoring applications (Osowski et al., 2004). However, model parameters in SVM (kernsels and parameters), which may be difficult to fine-tune in real-time systems, are crucial in obtaining good quantity estimation.

- **K-Nearest Neighbors (KNN):**

KNN is a non-parametric algorithm that categorizes the instances in a majority of the k closest training points in the feature space. It is simple, interpretable and effective in situations where the data distribution is well-behaved. In the field of PHM, KNN has been used in various activities like posture recognition, fall detection, and activity class. KNN is intuitive, but it is computationally expensive with large data sets and is affected by distance metric and the value of k (Altun et al. 2010). Moreover, its performance becomes worse in high dimensional spaces because of the curse of dimensionality.

- **Decision Trees (DT):**

Decision Trees are flowchart-like models that are used for both classification and regression. Typically, they are appreciated for interpretability and simplicity to be deployed in embedded systems. In PHM, DTs have been applied to predict sleep stages, diabetes risk and lifestyle behaviors assessment. Because of their hierarchical nature, they are well suited to the construction of if-then decision rules which are transparent and clinically useful. However, DTs are susceptible to overfitting especially if the trees become deep without pruning or regularization (Zhou et al., 2019).

- **Comparative Insights and Hybrid Usage:**

Each of these algorithms has its own unique strengths in that SVM is powerful for complex boundaries, KNN is easy to use and adaptable, and DTs provide transparent decision-making. In practice, PHM systems would often use a combination of these algorithms, along with feature selection algorithms, normalization algorithms, and ensemble learning algorithms (e.g., Random Forests, Boosted Trees) in order to improve the performance and generalizability of these systems. The algorithm selection is based on factors like the nature of the health condition being monitored, the volume and type of data, the need for real-time processing, and the level of interpretability required.

The use of ML algorithms in PHM continues to evolve with the development of optimized learning frameworks and automated model selection techniques. Nonetheless, understanding the strengths and limitations of each approach remains essential for effective algorithm selection and deployment in real-world healthcare applications.

### 2.3.2 Deep Learning Algorithms

Personal Health Monitoring (PHM) systems can benefit from Deep Learning (DL), a subfield of Machine Learning (ML), which has emerged as a revolutionary factor in the design and optimization of PHM systems. Unlike conventional traditional ML models, which generally rely on the manual extraction of features, DL architectures automatically acquire hierarchical features representations from raw data which lead to more accurate and more scalable solutions. In PHM, in which data is usually high dimensional, continuous, and multimodal (e.g., time series, image, and audio data), deep learning algorithms have demonstrated impressive performance in a number of applications.

Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs) and their extensions like Long Short Term Memory (LSTM) networks are among the top 5 DL models that find their use in PHM. Each architecture has its own merit in processing data related to health and their choice is usually based on the kind of the input signals and the application, which is required.

**1. Convolutional Neural Networks (CNNs):**

CNNs are very good at learning spatial features of structured data, and are therefore widely used for medical imaging and sensor classification. In PHM, CNNs are commonly used for interpretation of data coming from wearable devices such as electrocardiogram (ECG), photoplethysmogram (PPG), and activity sensors. For instance, one-dimensional CNNs have been successfully used on raw

ECG waveforms for arrhythmia detection, which outperforms traditional classifiers (based on ML) in sensitivity and specificity (Rajpurkar et al, 2017). CNNs are also used for emotion recognition from physiological signals, sleep stage classification from multi-channel biosignals and so on. The advantages of using them for embedded PHM applications are their ability to capture localized patterns, dimensionality reduction through pooling and fast deployment. However, CNNs do not handle modeling sequential dependency or temporal dynamics which are often important in longitudinal health monitoring. This limitation has resulted in the adoption of RNN-based models in a growing number of PHM contexts where time series analysis is of central importance.

**2. Recurrent Neural Networks (RNNs):**

RNNs can be used for the sequential data processing by keeping the internal memory states, and are well suited for time-series data analyses in PHM systems. They are commonly used with data with continuous monitoring of physiological parameters such as cardiac activity or respiratory rate or glucose levels. The basic RNNs have been shown to be useful in early event detection and trend analysis, but these RNNs suffer from the vanishing gradient problems when handling long sequence streams. To overcome these limitations, more complex variants of RNN have been used, namely LSTMs and GRUs which have superior retention of long-term dependencies and more stable training (Lipton et al., 2016).

**3. Long Short-Term Memory Networks (LSTM):**

LSTM network is the most representative RNN network applied to PHM because of their success in modeling complex temporal relationships. They contain gated mechanisms (input, forget and output gates) which allow them to selectively keep or forget information across time steps in order to give more accurate forecasting and anomaly detection.

LSTM-based models have been used in applications such as:

- Predicting future blood glucose levels in diabetic patients based on prior readings and lifestyle variables.
- Monitoring mental health by analyzing mood and activity logs over time.
- Forecasting epileptic seizures using EEG time-series data.

LSTM networks are also increasingly combined with CNNs in hybrid architectures, allowing the model to extract both spatial and temporal features (e.g., CNN-LSTM models for multi-sensor fusion in fall detection or stress recognition).

**Challenges and Optimization Techniques:**

Despite their strengths, deep learning models require large volumes of labeled data, significant computational resources, and careful hyperparameter tuning. These problems can become especially daunting in the PHM domain, where the availability of annotated datasets is limited or there is a need to use low-resource wearable devices. To prevent these limitations techniques such as

transfer learning, data augmentation, and model compression (e.g., pruning, quantization) are often employed.

Furthermore, for interpretability is always a concern. Deep models can be considered as "black boxes" and this could impede their occurrence in clinical settings. As a consequence, explainable AI (XAI) frameworks are being coupled with DL-based PHM systems in order to bring transparency to the decision-making processes in order to build trust and to comply with regulatory requirements. Deep learning algorithms have been commonly used to greatly extend analysis capability of PHM systems with high accuracy in complex tasks such as pattern recognition, temporal prediction, and multi-modal data fusion. Their deployment is still increasing, aided by the development of wearable computing, edge AI, and federated learning. However, the future of these kind of models cemented by continuous efforts to achieve an accurate, efficient, and interpretable model within the confines of the real-world healthcare setting.

### 2.3.3 Ensemble and Hybrid Models

As Personal Health Monitoring (PHM) systems evolve to address increasingly Once PHM systems developed a system to meet complex healthcare needs, the drawbacks of relying on a single algorithm became obvious. As a result, researchers and developers are looking to ensemble and hybrid modeling approaches for improved prediction accuracy, robustness and generalization capability. These approaches use a combination of algorithms, either of the same kind or from different categories, to overcome the weaknesses of individual models and to exploit complementary strengths.

*Figure 3* illustrates a typical architecture of an ensemble and hybrid model used in PHM systems. In this architecture, input data such as physiological signals from wearable sensors are simultaneously fed into multiple models:

- CNN (Convolutional Neural Network): Extracts spatial features, effective for signals like ECG or PPG.
- LSTM (Long Short-Term Memory): Captures temporal dependencies in sequential data such as heart rate variability or glucose level trends.
- Performing classification on the basis of extracted features is also another method of classification, functioning using SVM (Support Vector Machine) which is valued for its robustness and interpretability. Each model works on the data separately and their outputs are combined in a fusion layer by majority voting, weighted averaging or meta-learning. This ensemble-hybrid approach gives a more generalized and accurate output, which may be a health alert, diagnosis, or recommendation.

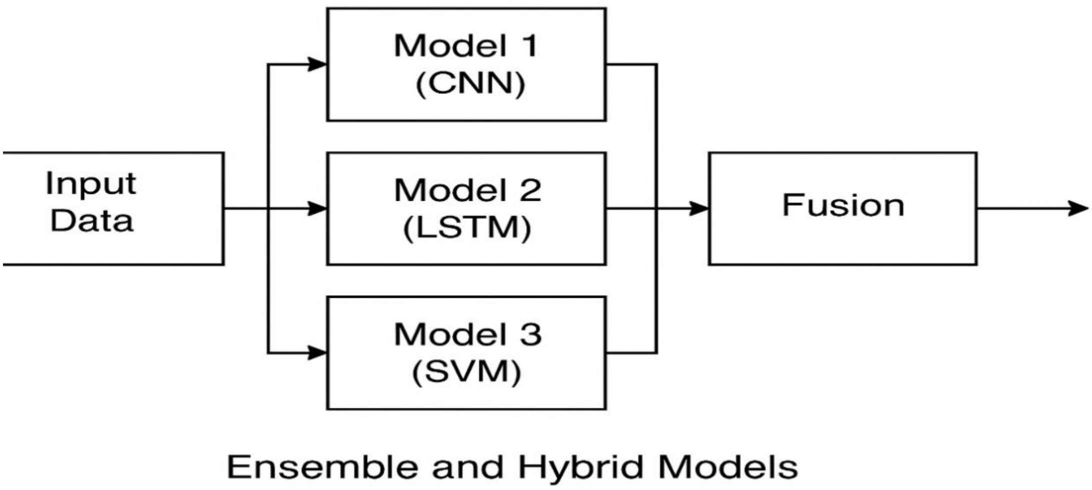


**Figure 3: Ensemble and Hybrid Models in Personal Health Monitoring Systems**

**Source: Author (2025).**

**1. Ensemble Models in PHM**

The techniques that combine the predictions of several base models to make a final, usually more accurate, prediction are known as ensemble learning. The most common ensemble strategies are used in PHM are bagging, boosting, and stacking.

Most bags have models that are classifiers.<|human|>Most bagging (Bootstrap Aggregating) has classifier models.

To minimize variance and avoid overfitting, bagging is applied in PHM to train several models (e.g. decision trees) on dissimilar random subsets of the dataset. Among the most popular approaches to bagging, the Random Forests have found notable applications in the sphere of activity recognition, falls detection, and cardiovascular-related risk prediction (Chen et al., 2017). They ensure that they are also appealing in real time, sensor based applications because of their capability to process noisy data and rank features in terms of importance.

• Boosting:

AdaBoost algorithms, Gradient Boosting Machines (GBM) and other types of boosting algorithms are trained successively such that each incorrectly learner addresses the errors made by the previous learner. Boosting has been used in PHM to detect stress, predict glycemic events and classify sleep stages. These models can usually be useful in imbalanced or more complicated datasets, but they are noise-sensitive and often need sensitive regularization (Zhou et al., 2021).

• Stacking:

Stacking is an addition of two or more heterogeneous base models (e.g., SVM, KNN, neural networks) and submits the predictions made by them to a meta-model to make the final decision. The approach has demonstrated potential in emotion recognition by wearables and multimodal data fusion, in which a variety of data types and patterns require various modeling capabilities (Wang et al., 2019).

**2. Hybrid Models in PHM**

The concept of hybrid models is an extension of ensemble learning in that various algorithmic paradigms are joined, commonly a combination of machine learning (ML) and deep learning (DL), or a combination of rule-based systems and data driven models. It is aimed at developing architectures that are capable of managing the temporal, spatial, and semantic complexity of health data in a more efficient way.

• ML-DL Hybrids:

One type of solution is to apply ML classifiers (e.g. SVM, random forest) to final decision-making, and deep networks (e.g. CNN or LSTM) to extract features. As an example, in ECG classification, CNNs can be used to obtain strong spatial features that can be further refined by a gradient boosting classifier to enhance interpretability and accuracy.

**2. Hybrid Models in PHM**

- The hybrid models are not confined to ensemble learning, typically involving the integration of various algorithmic paradigms, frequently comprising machine learning (ML) and the deep learning (DL), or rule-based systems and data-driven models. The idea is to develop architectures that are capable of managing the temporal, spatial, and semantic complexity of health data in a better way.
- **ML-DL Hybrids:**
- There is also a popular practice of determining the final decision with the help of ML classifiers (e.g., SVM, Random Forest), and feature acquisition with the help of deep networks (e.g., CNN or LSTM). As an example, when used in ECG classification problems, CNNs can be used to generate strong spatial representation that is subsequently feed into a gradient boosting classifier that is more interpretable and accurate.
- **Rule-Based + Learning-Based Integration:**
- But there are also hybrid systems that combine domain knowledge in the form of expert defined rules with adaptive learning algorithms. It is especially applicable to clinical PHM contexts where critical decisions that are safety-related need some transparency and validation. As an example, pre-defined clinical thresholds can be used to raise an alarm, and AI models can also give probability scores to prioritize based on the context (Ting et al.
- **Multimodal Hybrid Architectures:**

In practice PHM, data may be provided by various sources (e.g., accelerometers, ECG, audio and user-reported symptoms). This data can be successfully integrated by hybrid models based on a mixture of parallel CNN-LSTM architectures, attention, and decision-level fusion approaches. These are robust architectures that have been optimized to provide better generalization in the performance among different user profiles and environments.

**Advantages and Considerations:**

The main benefit of ensemble and hybrid models is that they have a better predictive performance and can be adapted to complex PHM tasks. They minimize model bias and variance, become more resistant to noise, and can also tend to offer more successful generalization to various datasets. These advantages however have trade-offs. Ensemble and hybrid systems are more computationally intensive, less interpretable and difficult to implement on resource-constrained devices like wearables. Furthermore, their effectiveness is strongly determined by the correct choice of models, training calibration, and preprocessing data strategies. With the increased prevalence of PHM systems and the resulting data, ensemble and hybrid solutions are likely to assume a more important role in the creation of robust, scalable, and clinically meaningful health monitoring solutions.

### 2.3.4 Comparison of Algorithm Effectiveness and Use Cases

Evaluating the effectiveness of artificial intelligence (AI) algorithms used in Personal Health Monitoring (PHM) systems requires a multifaceted approach that considers accuracy, computational efficiency, interpretability, scalability, and suitability for real-time deployment. No single algorithm universally outperforms others across all health monitoring tasks; rather, each algorithm demonstrates strengths and trade-offs depending on the nature of the health condition, the type and quality of input data, and the target population.

This section compares the most widely used algorithms—Machine Learning (ML), Deep Learning (DL), and Ensemble/Hybrid models—across various PHM use cases, drawing on performance metrics and practical deployment considerations.

**1. Accuracy and Predictive Power**

- **Deep learning algorithms**, particularly Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, are always at the top of the list when it comes to tasks that require complicated temporal or spatially looking data, unlike ECG classification, activity recognition, or sleep stage detection. F1-scores of over 90% in arrhythmia detection have been reported when CNNs are used (Rajpurkar et al., 2017).
- **Machine learning models** Similarly to the Support Vector Machines (SVM) and Random Forests (RF) will frequently outperform deep models on structured data whose characteristics are known, e.g. physical activity classification or chronic disease prediction (Altun et al., 2010).
    - **Ensemble methods** Random forest and Gradient Boosting Machines (GBM) among others often outperform single ML predictors in terms of taming variance and creating a better generalization. GBM models were found to be more robust in various test sets in stress detection using wearable data, e.g., GBM models (Zhou et al., 2021).

**2. Interpretability and Clinical Trust**

- **Decision Trees and SVMs** are more interpretable than deep models, making them preferable in contexts where explainability is crucial—such as clinical decision support or regulatory approval.
- **DL models**Although strong, they are usually criticized as black boxes. This is a challenge in the healthcare sector where trust and traceability are critical. Recent analyzeable AI (XAI) endeavors attempt to fill this void via ideas, such as saliency maps and attention..
    - **Hybrid approaches** that explicitly combine deep feature representation with interpretable classifiers (e.g. CNN + SVM) are becoming popular in PHM in order to achieve both accuracy and transparency.

**3. Computational Efficiency and Resource Demands**

- **ML models** (e.g., KNN, DT) tend to use less computational power and memory, which is endorsed to utilize in edge deployment on wearables or on a mobile system with scarce resources.
- **DL models** demand significant processing capacity and often require GPUs or specialized hardware. However, advancements in edge AI and model compression (e.g., pruning, quantization) are enabling more efficient deployment of deep models in embedded environments (Sze et al., 2017).
- **Ensemble models** may suffer from higher inference time and complexity, depending on the number and type of base learners. They are typically deployed in cloud-based PHM architectures where latency is less critical.

### 4. Use Case Suitability

**Table 1: Comparison of algorithms by PHM use case.**

| Use Case | Recommended Algorithms | Justification |
|---|---|---|
| ECG classification | CNN, LSTM, CNN+SVM | High-dimensional time-series data; benefit from deep feature extraction |
| Activity and posture recognition | Random Forest, KNN, SVM | Structured sensor data; high interpretability and computational efficiency |
| Sleep stage detection | LSTM, CNN, GBM | Temporal pattern recognition; need for high sensitivity and adaptability |
| Stress and emotion monitoring | Ensemble models (GBM, stacking), Hybrid CNN-LSTM | Multimodal inputs; benefit from robustness and data fusion |
| Fall detection | DT, SVM, CNN | Real-time decision-making; balance between speed and accuracy |
| Diabetes/glucose trend prediction | LSTM, Hybrid DL-ML | Long-range dependencies in metabolic patterns |

### 5. Generalization and Scalability

- **Ensemble and hybrid models** provide better generalization ability by learning different patterns from different domains. They are especially appropriate in case of heterogeneous or imbalanced training data.
- **Deep learning models**, with enough training data, scalable across a range of different health applications However, they may struggle when we have small datasets or they are moved to new populations without being retrained.
- **ML models** are easier to calibrate and deploy and may not work easily with high-dimensional or multimodal input without the help of careful feature engineering
- The relative performance and suitability of AI algorithms in PHM would not only be determined by raw performance indicators but would also depend on context-specific considerations such as availability of data, need for interpretability, real-time limitations and deployment environment. Whereas deep learning models continue to dominate in terms of predictive accuracy for more complicated tasks, traditional ML models continue to be relevant (due to their relative simplicity and transparency). Different ensemble and hybrid strategies still promise to be a good compromise between them by combining the best from many different paradigms.

- In future PHM systems it will be likely that adaptive context-aware algorithm selection will be deployed to efficiently and practically align computational models with user's profile and aim of monitoring, maximizing both efficiency and clinical value.

## 2.4 Technical and Ethical Challenges

### 2.4.1 Technical Limitations in AI-based PHM Systems

Despite the potential progress in the integration of Artificial Intelligence (AI) into PHM systems, there are still many technical limitations that prevent their performance, scalability, and adoption in real-world healthcare. These issues range from data-related challenges, model design and deployment limitations, and systems integration, and hinder the development of reliable, fair, and user-centric PHM technologies.

**1. Data Quality and Heterogeneity**

One of the most significant constraints in PHM based on AI is the inconsistency of data quality. The data collected by wearable devices and mobile health applications are noisy, incomplete, or inconsistent because of hardware constraints, noncompliance of users, or environmental interference (Rehman et al., 2020). Variability in the precision of sensors from different device manufacturers add another layer of complexity which make the trained models less generalizible.

Additionally, PHM data is very heterogeneous, including the different formats (time series signals, text entries, audio inputs, and contextual metadata). This multimodal character makes the transformation of data preprocessing and feature extraction a technically-demanding process with the need for sophisticated fusion and alignment techniques.

**2. Limited Labeled Datasets and Imbalanced Classes**

AI algorithms, especially algorithms that use supervised learning methods, need vast amounts of labels before they can function effectively. In the healthcare domain, high-quality labelled datasets are hard to acquire due to privacy issues, restricted availability of clinical annotation, and the expensive cost of manual labelling of a subject. However, most PHM applications are affected by small sample sizes and class imbalance, i.e., some conditions (e.g., rare cardiac arrhythmias or epileptic events) are underrepresented. Imbalanced datasets have been linked to biased classifications with good performance on majority classes but poor assessment of rare but clinically important events. He or she or it is true that there are techniques available for building models that are centered on resampling, synthetic data generation (e.g., SMOTE), and cost-sensitive approaches that come with a trade-off between model stability and accuracy.

**3. Real-Time Constraints and Computational Limitations**

The majority of PHM applications require the real-time or near real-time processing of the data, especially for safety-critical applications like fall detection or arrhythmia monitoring. However, the computation complexity of the state of the art AI models, especially deep learning architectures, may lead to high latency, energy usage and memory consumption, which is unsuited with the hardware limitations of edge devices (Sze et al., 2017).

While there are methods that have attempted to alleviate some of this, namely edge AI and model optimization methods (e.g. pruning, quantization), it is still a significant challenge to deploy complex AI models on lightweight devices without sacrificing performance.

**4. Lack of Standardization and Interoperability**

Universal standards for data formats, communication protocols and model evaluation metrics are lacking in PHM systems, which hinders the integration and scalability of the system. This lack of standardization is making it difficult to deploy AI models on different platforms and environments, which results in siloed systems with limited interoperability [Shickel et al. 2018]. Moreover, different evaluation metrics are used in different studies, i.e. accuracy, AUC, precision-recall, it is difficult to compare the performance of the models and to replicate the results.

**5. Model Interpretability and Clinical Integration**

AI models, especially deep learning models, have been criticized for being "black-boxes," where internal decision-making processes are not transparent to users and clinicians. This lack of transparency makes it impossible for there to be clinical trust and regulatory approval of AI-based PHM devices. Although the field of Explainable AI (XAI) is advancing, the application of interpretability tools in PHM systems that are technically effective and clinically meaningful is still limited (Tjoa & Guan, 2020). Furthermore, the aim of contextualizing AI outputs with clinical workflow to integrate or incorporate them with electronic health records (EHRs) poses a challenge that is currently beyond the reach of most systems.

**6. Model Robustness and Generalizability**

AI models that are trained with controlled datasets sometimes fail to perform under more natural variations, for example changes in the position of a sensor, or user's behaviour, or even the environmental conditions. Robustness to such variations is the key to PHM systems to be reliable over time and in varied user populations. However, many current models are not carefully validated across demographic groups or device settings, raising concerns about fairness and equity of algorithmic performance.

By addressing the technical limitations of AI in PHM systems, it is possible to ensure their reliability, safety, and scalability. Future studies should focus on building large data processing pipelines, making models more generalizable through domain adaptation, real-time edge computing, and the development of the adoption of interoperability standards. Further, the call for increasing model transparency and user trust will play a crucial role in the translation of the promising AI innovation in PHM to clinically valid and ethically sound solutions.

### 2.4.2 Ethical Concerns: Privacy, Bias, and Transparency

As artificial intelligence (AI) continues to increase its role in Personal Health Monitoring (PHM) systems, a slew of ethical issues in areas of data privacy, algorithmic bias and model transparency have been identified. These issues have real-world implications for the trust of users, healthcare equity and social acceptability of the use of AI-powered technologies. Thus, ethical robustness is an important requirement for the responsible development, implementation, and governance of PHM systems.

**1. Data Privacy and Informed Consent**

PHM systems use constant and often passive data acquisition from wearable sensors, mobile applications, and the cloud platform. This data contains very sensitive information on the individual including heart rate patterns, sleeping bellows, location data and even emotional states. The continuous nature of the monitoring process increases risks associated with unauthorized access, data breaches, and surveillance (Hagendorff, 2020). One of the ethical concerns is the sufficiency of informed consent. Most users have no idea how much information is being collected about them, analyzed and possibly divulged to third parties. Furthermore, long and complicated privacy policies can make the consequences of data use unclear. Ethical

PHM design must includes transparent consent mechanisms, user-controlled data sharing settings, and regulatory adherence, such as the General Data Protection Regulation (GDPR) and the Health Insurance Portability and Accountability Act (HIPAA). Novel solutions such as federated learning and differential privacy are emerging to enable collective intelligence of distributed systems without centralizing raw data in order to maintain user confidentiality while training artificial intelligence models.

**2. Algorithmic Bias and Fairness**

AI models used in PHM are highly sensitive to the quality and diversity of the data on which they are trained. If training datasets are skewed toward certain demographic groups—based on age, gender, race, or socioeconomic status—the resulting models may yield biased outputs. For example, a fall detection system trained predominantly on young male subjects may underperform when deployed with elderly females, increasing the risk of false negatives or inappropriate alerts (Obermeyer et al., 2019).

Such biases can perpetuate or even exacerbate existing health disparities. In the context of PHM, where real-time feedback informs critical health decisions, algorithmic fairness is not just a technical goal but a moral imperative. Developers must proactively audit datasets for representation, employ fairness-aware learning algorithms, and conduct stratified performance evaluations to detect and mitigate bias across different subgroups.

**3. Transparency and Explainability**

Model transparency (i.e., the ability to understand how an AI system comes to its decisions) is an important foundation of ethical requirements, especially in healthcare settings where accountability is key. Most deep learning models applied to PHM (e.g., CNNs, LSTMs) are highly complex and low-interpretable, which makes it challenging for users, clinicians, or regulators to trust or validate their results (Tjoa & Guan, 2020). Explainable AI (XAI) approaches are increasingly being applied for PHM systems in order to address this gap. Model interpretability tools like SHAP (SHapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations) are used to determine the contribution of features to a specific prediction. However, the practical application of such tools is limited in many commercial products in the field of PHM, and not many provide easy-to-use visualizations of model reasoning. Transparency is also associated with user autonomy. In line with the principles of ethical AI, health monitoring systems need the capability for users to contest, challenge or choose not to act on AI-generated recommendations, especially in cases where such recommendations impact health-relevant decisions.

**4. Data Ownership and Control**

The question of data ownership is becoming a hot topic in ethical debates. If PHM systems capture the data from users, ownership and management is still usually with technology providers or third-party analytics companies. This asymmetric information results in power imbalances and may be used for secondary purposes of health data (e.g. marketing, insurance profiling) that users do not expect or that may conflict with public interest. This is why advocacy for data sovereignty models - a model where individuals are the data owners and they can control how and when their data is used, particularly as it relates to their health - is on the rise. Ethical guidelines should be enforced for PHM platforms to follow open policies on data access, retention and deletion to allow users to have a meaningful control over their personal health information. The ethical issues in AI-based PHM systems are multifaceted, intertwined and extremely impactful. Societies are facing them in an interdisciplinary way, involving technological innovation but also legal protection, a user centered design and a normative reflection. Ensuring privacy, reducing bias, and increasing transparency are not only ethical requirements, but are essential requirements for building trust, compliance, and equitable results for digital health monitoring.

### 2.4.3 Legal and Regulatory Issues in AI-Powered Monitoring

The fast uptake of Artificial Intelligence (AI) in Personal Health Monitoring (PHM) systems opens up a complex and changing landscape of legal and regulatory challenges. Such issues are not just from a safety and efficacy perspective but also to the basic issue of data protection, liability, accountability and national/international legal compliance of AI-enabled devices. These matters must be dealt with for the responsible innovation of PHM devices, for the protection of the rights of patients and for the credibility of PHM technologies in the eyes of the public.

**1. Medical Device Regulation and AI Classification**

One of the main legal issues with AI-enabled PHM systems is the nature and extent of their subjecting to existing medical device legislation. In the U.S. and the European Union, software applications that provide diagnostic or therapeutic functions are considered a medical device and regulated by the regulatory agency in each region.

In the US, the Food and Drug Administration (FDA) uses a risk-based approach to the oversight of AI software, as part of its Digital Health Framework. Depending on the risk class and intended use, AI-enabled PHM applications made for medical applications (e.g., arrhythmia detection, seizure prediction) may, dependent upon their intended use, be required to be cleared for premarket approval or approved (FDA, 2021).

In the EU, the scope of applicability of the Medical Device Regulation (MDR) has already been expanded to encompass both standalone software and artificial intelligence-based tools. Medical devices are PHM systems which need to be shown to be conformant with safety, performance and cybersecurity requirements.

One of the main regulatory difficulties is that AI algorithms are usually adaptive by nature and can change by ongoing learning after implementation. It is also shown that current regulatory systems are not well suited for such dynamic updates; agencies such as FDA are starting to investigate "Software as a Medical Device (SaMD)" constructs in which measures of real-time tracking and post-market surveillance are embedded into the systems..

**2. Data Protection and Cross-Border Data Flow**

AI-based PHM systems also process large amounts of personal and sensitive health information on a regular basis and generate important legal questions regarding data ownership, consent and international data transfers. Key legislation for this field is: Valid reasons by researchers for not seeking informed consent from participants may include: - Suspecting a research subject of fraud. - Migratory research subjects. - Researchers who are ignorant of a subject's ethnicity at the time of consent. - When seeking informed consent with potential participants, researchers must use language that is easy for them to understand. - Participating in research that tests a medical device or procedure. - Being language deficient. - For example, since many patients would not want to disclose their ethnicity to a participant, researchers might be compelled to continue obtaining consent to participate in research. - Evidence of a potential risk of harm to participants from Health Information Technology Act (HIT), also known as Health Information Technology (HIT) in the U.S. or Health Information Technology Staff is responsible for regulating the use and distribution of protected health information (PHI) by covered entities and business associates in the

U.S. Legal issues become more complex when PHM data is stored or processed in cloud platforms that span borders. The absence of international standards on data sovereignty makes it difficult for developers and healthcare providers to comply. Mechanisms like Standard Contractual Clauses (SCCs) and Binding Corporate Rules (BCRs) which are most commonly needed to legalise data transfers under GDPR have an added layer of legislative complexity and administrative overhead. (Tschider, C. A., 2021)

3. **Liability and Accountability in Automated Decision-Making**

The question of liability when an AI-based PHM system fails - either by giving wrong advice, failing to detect a critical event or causing harm through system error - is a significant open legal question. Traditional tort law imposes liability on the basis of negligence or product defect, but it is not so simple to apply these concepts to self-learning algorithms.

• Who is liable? The developer, the device manufacturer, the healthcare provider or the end-user?

• What constitutes a defect? Is it a data problem (bad training data), a programming problem (bad algorithm) or a problem with no humans in the loop to learn to act appropriately? Algorithmic accountability is a relatively new concept that legal scholars and regulators are currently discussing which could include things like being required to document model logic, audit trails, and explainability requirements. Some proposals call for the establishment of special AI insurance schemes or models of liability on a shared basis for medical applications with a high risk of misconduct.

4. **Certification and Standardization Gaps**

One of the regulatory gaps is that there are no globally agreed standards for certifying AI algorithms for use in healthcare. While there are some voluntary standards (e.g. ISO/IEC 62304 for medical software lifecycle, ISO/IEC 27001 for data security), there is no global certification protocol for AI models used in PHM.

The absence of any regularization has also led to greater fragmentation, where different countries or governance bodies use different frameworks. It also causes uncertainty for developers who want to get their products approved on the market in multiple jurisdictions. Efforts to harmonise regulatory principles like the International Medical Device Regulators Forum (IMDRF) have gone some way in establishing harmonised regulatory principles but implementation remains low.

AI-driven PHM systems will push the potential of legal and regulatory frameworks to the limits in an unparalleled way. Technological development, health care and legal professionals and regulators must collaborate to ensure compliance. Future activities need to be focussed on dynamic regulatory models that allow for learning continuously, crossing borders on data governance, accountability of algorithms, and clear certification routes. Without such underpinnings, PHM technologies can no longer be long lasting or be used in an ethical manner.

## 2.5 Review of Relevant Previous Studies

A vast literature has been established in recent years regarding the use of Artificial Intelligence (AI) in Personal Health Monitoring (PHM). Its diverse research ranges across several subdomains including chronic disease management, remote monitoring, wearable technology and deep learning diagnostics. This paper provides a synthesis of the significant research that has been conducted, with the aim of ascertaining the predominant directions of research, methodological patterns, and the emergence of a developing agenda of AI-enabled PHM systems.

1. **Pranam et al. (2024):**

The present study examined the incorporation of artificial intelligence methods such as machine learning, natural language processing and computer vision in PHM systems for smart homes. The authors underscored the move from reactive health interventions to proactive (context-aware) environments. Their results demonstrated not only the massive potential of AI technology to improve both safety and individualization through real-time adaptation to users' behavior patterns. However, the research also elucidated the importance of addressing user centric problems like availability and customization for system adoption and long term utility (Pranam et al., 2024).

2. **Palakurti (2023):**

Palakurti made a comprehensive review of trends in wearable health monitoring technologies focusing especially on artificial intelligence (AI) powered real-time health management. The study highlighted the increasing relevance of wearable sensors and mobile to enable continuous monitoring but also found issues with algorithm optimization, user trust and hardware limitation. While acknowledging how promising the performance of AI is in enabling early detection and tracking of behavior, it was concluded that usability and system responsiveness remain significant barriers against their widespread adoption (Palakurti, 2023).//////////

**3. Medjaher & Tran (2023):**

In a broad systematic review, Medjaher and Tran analyzed the landscape of AI applications across diverse PHM use cases. They categorized AI implementations into three major functions: disease prediction, behavioral tracking, and health trend analysis. Their synthesis revealed the predictive strength of AI algorithms in identifying risk patterns and anomalies; however, they also reported recurring concerns about data privacy, interoperability between systems, and the scalability of AI solutions in real-world deployments. The study called for more standardized evaluation metrics and regulatory frameworks to ensure ethical integration of AI in health monitoring contexts (Medjaher & Tran, 2023).

**4. Gulshan et al. (2016):**

This landmark study explored the diagnostic capacity of convolutional neural networks (CNNs) in the context of diabetic retinopathy screening. By utilizing a dataset of over 120,000 retinal fundus images, the researchers demonstrated that the AI model achieved a diagnostic accuracy comparable to—or in some cases exceeding—that of board-certified ophthalmologists. The study was among the first to validate deep learning performance against clinical gold standards, paving the way for future AI-assisted screening tools in PHM and telemedicine settings (Gulshan et al., 2016).

**5. Dai et al. (2023):**

Extending the work of Gulshan et al., this study applied deep learning models to a broader spectrum of ophthalmic diseases, leveraging a dataset of more than 210,000 annotated images. The performance of AI was compared against both non-physician graders and experienced clinicians, with results confirming the model's reliability across multiple diagnostic categories. The study contributed valuable insights into the feasibility of integrating AI-based diagnostic tools into population-level health monitoring and highlighted the potential for AI to augment decision-making in resource-constrained environments (Dai et al., 2023).

**6. Shaik et al. (2023):**

Shaik et al. conducted a comprehensive review of artificial intelligence in remote patient monitoring systems. They analyzed how AI-powered architectures—built on wearable IoT sensors, edge/cloud computing, and federated learning—facilitate early detection of patient deterioration, personalize health parameter monitoring, and learn behavior patterns via reinforcement learning. The study discussed key challenges such as data privacy, real-time processing constraints, and integration across heterogeneous platforms (Shaik et al., 2023).

**7. Nie et al. (2025):**

Nie et al. proposed a novel deep learning framework to infer “AI-derived vascular age” from photoplethysmography (PPG) signals using data from nearly 212,000 individuals. Their model demonstrated that a vascular-age gap exceeding nine years strongly correlated with higher risk of cardiovascular events—including diabetes, hypertension, and mortality—thus establishing a new digital biomarker for population-level monitoring (Nie et al., 2025)

**8. Ye et al. (2024):**

Ye et al. presented a dynamic, activity-aware health monitoring system (DActAHM) powered by deep reinforcement learning and SlowFast models. The system adapts monitoring frequency and content to the user's current physical activity, optimizing resource usage. Experimental results showed a performance gain of 27.3 % over static monitoring approaches, highlighting the value of contextually adaptive AI in real-time health tracking (Ye et al., 2024).

Collectively, these studies reflect a rapidly maturing field with several converging trends. First, there is growing emphasis on real-time, user-centric monitoring, facilitated by wearable devices and adaptive learning models. Second, chronic disease applications remain a dominant focus, due to their long-term data needs and the benefits of AI in trend analysis and risk prediction. Third, validation against clinical gold standards is becoming more prevalent, particularly in image-based diagnostics, highlighting AI’s role not just in monitoring but also in early detection.

Nonetheless, cross-cutting challenges persist. Issues of data privacy, regulatory compliance, algorithmic fairness, and user engagement remain under-addressed in many implementations. Moreover, few studies offer longitudinal validations or assessments of real-world deployment outcomes, limiting the ability to generalize findings across populations and settings.

The reviewed studies demonstrate the breadth and depth of current research on AI in PHM, showcasing the technology's potential in enhancing health outcomes through smarter, more responsive monitoring systems. However, the literature also reveals ongoing challenges that must be addressed to ensure safe, equitable, and scalable deployment. These include not only technical and clinical issues but also ethical, regulatory, and user-experience dimensions that future research must take into account.

# Chapter three
# Methodological Review and Comparative Analysis

## 3.1 Overview of the Review Approach

To provide a structured and comprehensive synthesis of existing literature on artificial intelligence algorithms applied in personal health monitoring, this study adopts a systematic review methodology. Systematic reviews are designed to minimize bias, enhance reproducibility, and allow for critical comparison across a large body of academic work. Given the fast-evolving nature of AI technologies and the increasing prevalence of wearable and mobile-based health systems, a structured synthesis becomes crucial to understanding which algorithms offer the most reliable, interpretable, and scalable solutions.

This chapter presents a detailed account of the methodology used to conduct the systematic review. It begins by outlining the review protocol, including the adoption of the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework to ensure transparency and methodological rigor. Inclusion and exclusion criteria were defined to focus the selection process on relevant, high-quality peer-reviewed research. In addition, multiple electronic databases were searched using carefully constructed Boolean queries to capture a broad and relevant set of studies.

Subsequently, data extraction was carried out to gather key information such as algorithm types, device categories, health domains, datasets used, performance metrics, and validation techniques. This information was synthesized into structured comparison tables, followed by a critical evaluation of trends, strengths, and limitations observed across the studies. Furthermore, quality assessment tools were applied to evaluate methodological robustness and reporting standards in the selected studies.

By combining a protocol-driven selection strategy with comparative analysis, this chapter lays the foundation for answering the research questions posed in Chapter One specifically, identifying which AI algorithms are most effective in PHM contexts, and how their performance varies across different applications and conditions.

### 3.1.1 Systematic Review Design and Justification

The adoption of a systematic review design in this study is grounded in the need for a rigorous, reproducible, and comprehensive synthesis of the growing body of research on artificial intelligence algorithms applied to personal health monitoring systems. A systematic review enables the identification, selection, and critical evaluation of studies using transparent methods, thereby minimizing bias and maximizing the validity and reliability of the findings. In the context of PHM, where technological advancements, algorithmic diversity, and heterogeneous data sources are rapidly expanding, such a review is essential to consolidate fragmented knowledge, identify trends, and reveal research gaps.

This review was conducted following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines, which provide a structured and standardized process for literature reviews. The choice of PRISMA is justified by its widespread adoption in biomedical and computational fields, as it ensures transparency in study identification, eligibility assessment, and inclusion procedures. The application of PRISMA is particularly relevant in this context due to the interdisciplinary nature of PHM, which encompasses computer science, biomedical engineering, and health informatics.

To ensure that the review captures the most relevant, recent, and high-quality research, a protocol was developed based on six foundational components:

1. **Clear Definition of Research Aim** : As described in Chapter One, the primary aim is to classify, compare, and evaluate AI algorithms used in PHM systems, with a particular focus on their performance, domain of application, and implementation context.
2. **Scope Specification**: Only studies that applied AI methods (including machine learning, deep learning, hybrid, or federated models) to PHM scenarios involving real-time or near-real-time monitoring using wearable or mobile-enabled devices were considered.
3. **Relevance Justification**: Given the proliferation of wearable technologies and the increased availability of health-related data, this review serves to synthesize the technical and practical insights derived from empirical applications.
4. **Methodological Rigor**: The design incorporates predefined inclusion and exclusion criteria (see Section 3.1.3), a multi-database search strategy (Section 3.1.4), and standardized data extraction protocols (Section 3.1.5).
5. **Quality Assurance**: To address variations in methodological strength, each selected study underwent a quality appraisal using standardized tools (see Section 3.1.6), ensuring that only studies meeting minimum scientific standards were retained.

6. **Comparative Intent**: Unlike narrative or scoping reviews, this review adopts a comparative orientation by organizing and evaluating studies along shared parameters (algorithm type, health domain, device modality, and performance metrics), thereby enabling the creation of performance comparison tables and classification schemas.

### 3.1.2 Review Protocol: PRISMA Framework and Adherence

To ensure transparency, replicability, and methodological rigor, this systematic review adhered to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines. PRISMA offers a standardized framework for reporting systematic reviews, facilitating clear documentation across all stages of the review process—namely, study identification, screening, eligibility assessment, and final inclusion. Its adoption in this review minimizes selection bias and enhances the reliability and reproducibility of the findings.

The review protocol was structured around four main PRISMA-compliant stages:

**1. Identification Phase**

A comprehensive literature search was carried out across five major academic databases:

- IEEE Xplore
- PubMed
- ScienceDirect
- SpringerLink
- ACM Digital Library

Boolean search strings were constructed using combinations of the following terms: ("personal health monitoring" OR "wearable health systems" OR "remote patient monitoring")

AND ("artificial intelligence" OR "machine learning" OR "deep learning" OR "federated learning") AND ("algorithm" OR "classification" OR "prediction")

This strategy yielded a total of 421 records, covering the publication period from January 2022 to March 2025.

**2. Screening Phase**

After removing 87 duplicate entries, a total of 334 unique records were screened based on their titles and abstracts. This stage focused on identifying studies relevant to the application of AI in PHM using wearable or mobile-based technologies. A total of 245 studies were excluded for reasons such as:

- Absence of AI-based methods
- Exclusive focus on hospital-based monitoring
- Being review articles, editorials, or conceptual papers without empirical evaluation

**3. Eligibility Phase**

The full texts of the remaining 89 studies were retrieved and assessed in detail against the predefined inclusion and exclusion criteria (as described in Section 3.1.3). An additional 74 studies were excluded at this stage for the following reasons:

- Lack of reported performance metrics
- Irrelevant application domains (e.g., non-health-related monitoring)
- Purely architectural studies without practical deployment in PHM scenarios

**4. Inclusion Phase**

Following the eligibility assessment, a final set of 15 studies was included in the systematic review. These studies met all inclusion criteria and provided sufficient methodological depth, performance reporting, and relevance to real-world PHM applications.

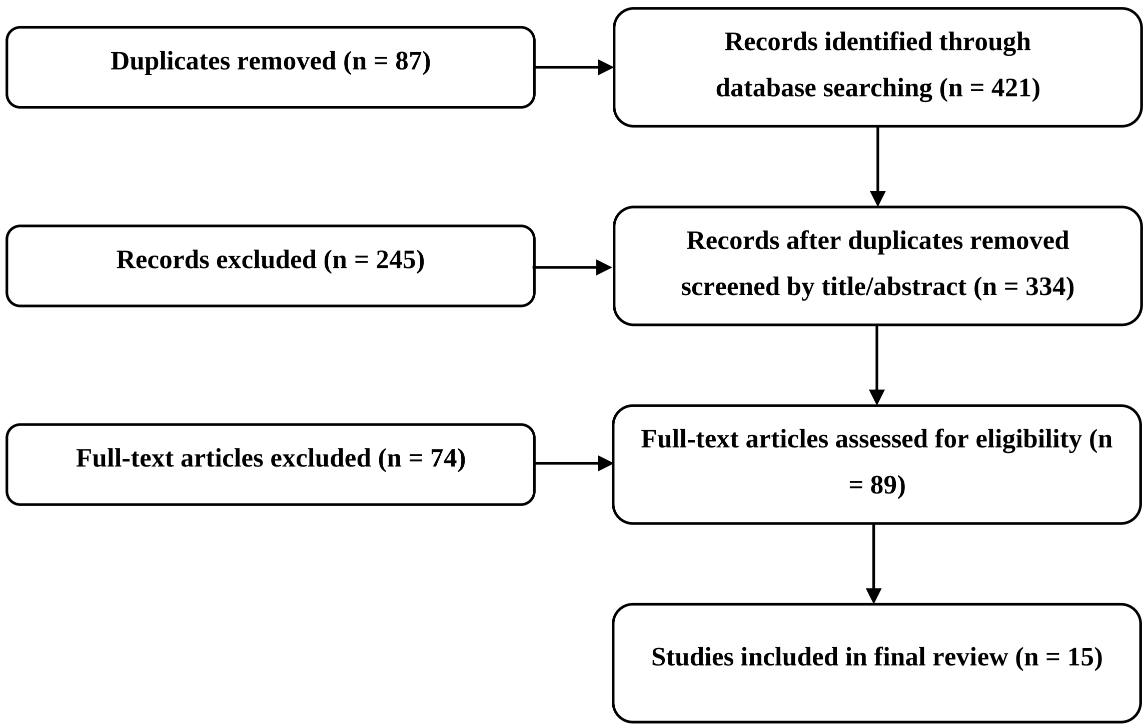


**Figure 4: PRISMA Flow Diagram for Study Selection**

**Source: Adapted from Page et al. (2021), PRISMA 2020 statement: an updated guideline for reporting systematic reviews, BMJ, 372:n71.**

To ensure the relevance, consistency, and scientific rigor of the included studies, a well-defined set of inclusion and exclusion criteria was applied during the screening and eligibility assessment phases of this systematic review. These criteria were formulated in alignment with the research objectives outlined in Chapter One and were specifically designed to target empirical studies that apply artificial intelligence (AI) algorithms within the context of personal health monitoring (PHM) using wearable or mobile-based technologies.

**Inclusion Criteria**

Studies were included in the review if they met all of the following conditions:

1. **Topical Relevance**:

The study should demonstrate the application of AI algorithms methods, or hybrid models—within PHM settings. These include real-time or continuous monitoring of physiological or behavioral parameters using digital health devices.

2. **Deployment in PHM Systems**:

Eligible studies must involve actual or simulated deployment of AI models integrated into wearable, mobile, or IoT-enabled health monitoring systems (e.g., ECG monitoring, sleep stage detection, blood pressure estimation).

3. **Reporting of Quantitative Performance Metrics**:

The study should report at least one recognized performance metric—such as accuracy, precision, recall, F1-score, AUC, MAE, or latency—to allow for valid comparison across models.

4. **Publication Type and Peer Review**:

Only original research articles published in peer-reviewed journals or high-quality conference proceedings between January 2022 and March 2025 were included.

5. **Language**:

   The article must be published in English to maintain consistency in interpretation and analysis.

6. **Accessibility**:

   The full text of the study must be accessible, either through open-access platforms or institutional databases, to enable in-depth review and data extraction.

**Exclusion Criteria**

Studies were excluded if they met any of the following conditions:

1. **Conceptual, Theoretical, or Review Papers**:

Articles lacking original empirical results, or limited to conceptual frameworks, taxonomies, or opinion pieces without algorithmic validation, were excluded.

2. **Hospital-Based or Non-Personal Monitoring Systems**:

Research focusing solely on clinical or hospital-based tools, or centralized monitoring systems not designed for individual use, was excluded.

3. **Lack of AI Implementation**:

Studies that relied solely on traditional statistical models, signal processing, or rule-based logic without incorporating AI algorithms were excluded.

4. **Absence of Device-Level Integration**:

Studies that only performed offline data analysis without integration into wearable or mobile systems were excluded—unless explicitly justified as directly transferable to PHM scenarios.

5. **Insufficient Performance Data**:

Articles lacking quantifiable performance metrics (e.g., accuracy, latency) or appropriate validation methods were excluded due to limitations in evaluability and comparison.

### 3.1.4 Study Search Strategy and Databases

The search strategy employed in this systematic review was designed to ensure comprehensive coverage, minimize the omission of relevant studies, and enhance replicability. A multi-phase search process was conducted across several academic databases known for their strong representation of computer science, biomedical engineering, and health informatics literature. This approach ensured balanced representation from both the computational and clinical aspects of personal health monitoring (PHM).

**Databases Selected**

Five prominent academic databases were used for the primary search phase:

1. **IEEE Xplore** – Focused on engineering and computing literature, particularly strong in wearable technologies, embedded systems, and AI algorithm design.

2. **PubMed** – A leading biomedical database indexing peer-reviewed medical and healthcare research.
3. **ScienceDirect (Elsevier)** – Covers a broad range of applied computing and biomedical engineering publications.
4. **SpringerLink** – Includes peer-reviewed journals and conference proceedings relevant to computational health sciences.
5. **ACM Digital Library** – Specializes in computer science research, with an emphasis on machine learning and mobile health applications.

These databases were selected for their complementary scopes and high relevance to the AI–PHM intersection.

**Search Query Formulation**

To maximize both precision and comprehensiveness, Boolean search queries were constructed across three thematic axes:

- **PHM Context**:

"Personal health monitoring" OR "remote patient monitoring" OR "wearable health system" OR "mobile health".

- **AI Techniques**:

"Artificial intelligence" OR "machine learning" OR "deep learning" OR "federated learning" OR "ensemble models".

- **Technical Implementation**:

"algorithm" OR "classification" OR "prediction" OR "real-time" OR "mobile application" OR "IoT".

The final search string implemented was:

("personal health monitoring" OR "remote patient monitoring" OR "wearable health system") AND ("artificial intelligence" OR "machine learning" OR "deep learning") AND ("algorithm" OR "classification" OR "prediction")

The query was slightly adapted to suit each database's syntax requirements (e.g., advanced filters in IEEE Xplore and PubMed).

**Search Limits and Time Frame**

- Time Frame: January 1, 2022 – March 15, 2025
- Language: English only
- Document Type: Peer-reviewed journal articles and full-length conference proceedings

A total of 421 records were retrieved initially. After removing 87 duplicates using automated tools like Zotero, 334 unique records remained for screening based on title and abstract.

**Supplementary Search Strategies**

To strengthen the comprehensiveness of the literature collection, the following supplementary strategies were employed:

- **Backward Citation Searching**: Reviewing reference lists of selected studies to identify additional relevant papers.
- **Manual Screening of Key Journals**: Reviewing recent issues of relevant journals such as *IEEE Journal of Biomedical and Health Informatics*, *Sensors (MDPI)*, and *Computers in Biology and Medicine*.
- **Selected ArXiv Preprints**: When technically robust and methodologically transparent, high-quality preprints from *arXiv.org* were included.

**Table 2: Summary of Database Search Results**

| Database | Records Retrieved | After Deduplication | Screened for Eligibility | Included in Final Review |
|---|---|---|---|---|
| IEEE Xplore | 128 | 103 | 27 | 5 |
| PubMed | 67 | 61 | 19 | 4 |
| ScienceDirect | 103 | 88 | 25 | 3 |
| SpringerLink | 62 | 51 | 15 | 2 |
| ACM Digital Lib. | 61 | 49 | 13 | 1 |
| Total | 421 | 334 | 89 | 15 |

As shown in Table 2, the distribution of retrieved and included studies reflects the breadth and filtering rigor of the search methodology applied in this review.

### 3.1.5 Data Extraction and Synthesis Process

Following the PRISMA-guided selection process, a rigorous and structured data extraction phase was implemented to facilitate subsequent comparative and thematic analysis.

**1. Data Extraction Procedure**

A standardized data extraction form was developed to collect essential variables from each of the 15 included studies. The form was designed to capture the following core attributes:

- **Study Identification**: Authors, title, publication year, journal/conference.

- **Health Domain**: The specific physiological or behavioral focus (e.g., cardiovascular health, mental health, neurological disorders, vital signs).
- **AI Algorithms Used**: Type of algorithm (e.g., machine learning, deep learning, federated learning, hybrid models), with specific methods noted (e.g., SVM, CNN, LSTM, XGBoost).
- **Device Type**: The hardware or system used for data acquisition, including wearable sensors, smartwatches, mobile apps, or microcontroller-based platforms.
- **Data Characteristics**: Data types (e.g., ECG, PPG, accelerometer, temperature), sources (public datasets or custom-collected), and sample sizes.
- **Performance Metrics**: Quantitative measures such as accuracy, F1-score, AUC, MAE, or system latency.
- **Validation Methods**: Whether the study used cross-validation, external testing, or real-time deployment.
- **Contributions and Limitations**: Highlights of the study's strengths, innovations, and reported challenges.

**2. Construction of the Master Summary Table**

A comprehensive summary table was constructed (see Table 3) to consolidate the extracted information and facilitate direct comparison across studies. This table serves as a foundational reference for subsequent analysis in Sections 3.2 and 3.3.

**Table 3: Summary of Selected Studies and Extracted Variables**

| Study ID | Year | Health Domain | Algorithm(s) Used | Device Type | Data Type | Accuracy / Metric | Validation Method |
|---|---|---|---|---|---|---|---|
| 1 | 2022 | Cardiovascular | CNN + LSTM + Semi-Supervised | Wearable ECG Patch | ECG | 90.2% | Comparison with baseline |
| 2 | 2023 | ECG Arrhythmia | CNN + Federated Learning + XAI | IoT-modeled ECG | ECG | 94.5–98.9% | 5-fold Cross-Validation |
| 3 | 2022 | Mental Health (Depression) | RF + XGBoost + SHAP | Wearable Wrist Sensor | HRV, EDA, Temp | 87.8–89.3% | 10-fold Cross-Validation |
| 4 | 2022 | Blood Pressure | SVR, KNN, ANN, GBR, RFR | Smartwatch (PPG) | PPG, SBP, DBP | MAE: 3.2–4.4 mmHg | 10-fold Cross-Validation |
| 5 | 2024 | Cardiovascular | CNN, RF, SVM, XGBoost, | Custom ECG Device | ECG | 92.6% | Internal Testing |

| Study ID | Year | Health Domain | Algorithm(s) Used | Device Type | Data Type | Accuracy / Metric | Validation Method |
|---|---|---|---|---|---|---|---|
| | | | Isolation Forest | | | | |
| 6 | 2023 | Neurological (Parkinson) | XGBoost, RF, SVM, KNN, DT, NB, GBM, LR | Wearable Accelerometer | Motion Sensor Signals | 94.5% | 10-fold Cross-Validation |
| 7 | 2025 | Behavioral (Smoking) | Federated CNN-LSTM + Secure Aggregation | Smartphone / Wearable | Motion + Contextual Behavioral Data | 95.3% | Federated Multi-Client Evaluation |
| 8 | 2025 | Sleep Monitoring | Transformer + Transfer Learning | Wearable (PPG + Respiration) | PPG, Respiratory Belt | 76.6% | Transfer Learning Validation |
| 9 | 2023 | Cardiovascular (Arrhythmia) | Residual CNN + Attention + SHAP | Wearable ECG Patch | ECG | 96.2% | Cross-validation + External Testing |
| 10 | 2023 | Vital Signs | Rule-based, Euclidean Scoring, Filters | Wearable + Mobile App | HR, $SpO_2$, Temp, GPS | 91% (Precision) | Live Pilot Test |
| 11 | 2023 | General Health | Fuzzy Logic, K-means, Lightweight DNN | Wearable + Smartphone | HR, $SpO_2$, Temp, Motion | 93.8% | Real-time Evaluation |
| 12 | 2023 | General Health | RF, DT, SVM, Naive Bayes | IoT-based Wearables | HR, Temp, $SpO_2$, Timestamp | 95.4% | Internal Validation |
| 13 | 2023 | Vital Signs | DNN, CNN, LSTM | Wearable + IoT | HR, Temp, BP, $SpO_2$, Motion | 97.2% | Stratified k-fold + Real Deployment |
| 14 | 2023 | Vital Signs | CNN (Edge-deployed) | IoT Wearable + ESP32 | Temp, HR, $SpO_2$ | 96.7% | Field Testing |
| 15 | 2022 | Vital Signs | XGBoost, RF, DT, GBM | Wearable + IoT + Cloud | HR, BP, $SpO_2$, Temp, Respiration | 96.2% | Cross-validation |

## 3. Data Synthesis Strategy

To transform the extracted data into meaningful insights, a multi-dimensional synthesis approach was applied:

- **Descriptive Synthesis**: Studies were grouped based on health domains (e.g., cardiovascular, neurological), types of AI models (e.g., classical ML vs DL vs hybrid), and device configurations.
- **Comparative Analysis**: Algorithmic performance was compared across studies using normalized metrics such as accuracy, F1-score, and latency. This allowed for the identification of high-performing models and contextual factors affecting accuracy.
- **Pattern Recognition**: Trends and recurring design choices were identified—for example, the prevalence of CNNs in signal-based monitoring, or the increasing use of federated learning in privacy-sensitive applications.
- **Thematic Coding**: Contributions and limitations noted in the studies were thematically coded (e.g., edge computing feasibility, lack of external validation, explainability challenges) to inform the challenges and recommendations in Chapter Four.

### 3.1.6 Study Quality Assessment

To ensure the methodological robustness and reliability of the included studies, a structured quality assessment process was carried out using standardized appraisal criteria. The goal of this process was to evaluate the scientific rigor, transparency, and internal validity of each study, thereby reinforcing the credibility of the synthesis and analysis presented in subsequent sections of this chapter. Quality assessment is particularly vital in systematic reviews involving algorithmic implementations, as variations in data quality, model validation, and reporting standards can significantly affect the comparability of results.

**1. Selection of Appraisal Framework**

Given the interdisciplinary nature of the included studies—spanning computer science, biomedical engineering, and health informatics—this review adapted a hybrid framework combining elements from two widely recognized tools:

- **CASP (Critical Appraisal Skills Programme)**: Originally developed for evaluating health and clinical studies, CASP was adapted to assess research clarity, objectives, methodological soundness, and result interpretation.
- **MMAT (Mixed Methods Appraisal Tool)**: Useful for interdisciplinary evaluations, MMAT elements were incorporated to assess methodological

appropriateness, data integrity, and internal consistency across quantitative designs.

To operationalize these tools in the AI and PHM context, a tailored checklist was developed that included the following seven criteria:

| No. | Quality Criterion |
|---|---|
| Q1 | Clear definition of study objectives and research questions |
| Q2 | Transparent description of AI model architecture and parameters |
| Q3 | Specification of data source, preprocessing, and labeling methods |
| Q4 | Use of appropriate validation techniques (e.g., cross-validation, external test) |
| Q5 | Reporting of performance metrics with comparative baseline or benchmarks |
| Q6 | Discussion of limitations, assumptions, and potential biases |
| Q7 | Relevance and applicability to real-world PHM scenarios |

Each study was scored on a binary scale: 1 = criterion met, 0 = not met or unclear. The maximum possible score was 7 points.

**2. Summary of Quality Assessment Results**

**Table 4: Quality Assessment of Included Studies Based on 7 Criteria**

| Study ID | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Total Score (out of 7) |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 7 |
| 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 7 |
| 3 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 6 |
| 4 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 7 |
| 5 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 4 |
| 6 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 7 |
| 7 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 7 |
| 8 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 6 |
| 9 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 7 |
| 10 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 3 |
| 11 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 7 |
| 12 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 6 |
| 13 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 7 |
| 14 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 7 |
| 15 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 7 |

Table 4 shows that 11 out of 15 studies scored the maximum of 7, indicating high methodological quality. Studies 5 and 10 scored lower due to limited reporting of validation techniques and incomplete discussion of limitations.

## 3.2 Summary of Selected Studies

This section presents a structured summary of the 15 studies selected for inclusion in this systematic review. Each study is categorized according to key attributes such as health domain, AI algorithm type, device used, data characteristics, and reported performance. The aim of this section is to offer a comprehensive descriptive overview that sets the stage for comparative evaluation and critical synthesis in the subsequent sections.

### 3.2.1 Studies on Cardiovascular and Arrhythmia Monitoring Using AI Techniques

**Study 1: Wang et al. (2022) – Hybrid Deep Learning for Real-Time ECG Analysis**

Wang et al. proposed a wearable ECG monitoring system that employs a hybrid deep learning model combining CNN and LSTM architectures for the detection of supraventricular premature beats and atrial fibrillation. The system utilizes a custom-designed wireless ECG patch capable of capturing time-series physiological data in real time. A semi-supervised learning approach was integrated to address noisy and incomplete labels. The model achieved an accuracy of 90.2%, significantly outperforming traditional baseline methods (58%). The study's main contributions lie in the integration of hardware-software co-design and its semi-supervised capability, aligning it directly with PHM objectives.

**Study 2: Raza et al. (2023) – Federated Learning with XAI for ECG Interpretation**

This study introduced a privacy-preserving ECG monitoring system based on federated learning and explainable AI (XAI). Data were sourced from the MIT-BIH Arrhythmia database, and the proposed framework incorporated a deep CNN autoencoder coupled with a classifier for arrhythmia detection. Notably, the federated setup enabled decentralized training, preserving patient data privacy. The model achieved an accuracy of 98.9% on clean data and 94.5% on noisy data, validated using 5-fold cross-validation. Additionally, an XAI module was employed to provide model interpretability, reinforcing clinical relevance and trustworthiness.

**Study 3: Alimbayeva et al. (2024) – Multimodel Comparison on a Custom ECG Platform**

This research involved the development of a real-time, custom wearable ECG device for early cardiovascular disease detection. A diverse set of machine learning and deep learning models was applied, including CNN, SVM, Random Forest, and Isolation Forest. Among these, CNN achieved the highest accuracy (92.6%). While the device design and edge integration were commendable, the study lacked external validation and comprehensive evaluation metrics, limiting its generalizability. Nonetheless, it demonstrates the feasibility of edge-AI deployment in PHM scenarios.

**Study 4: Woo et al. (2023) – Interpretable DL for Arrhythmia via Residual CNN and SHAP**

Woo et al. presented an interpretable deep learning model for arrhythmia detection using single-lead ECG signals. The system utilized residual convolutional layers with attention mechanisms and SHAP-based explainability tools. Real-world wearable ECG data and public datasets were used for training and validation. The model achieved an overall accuracy of 96.2% and an F1-score of 0.942 for atrial fibrillation. The inclusion of interpretability made this study particularly relevant for clinical adoption and physician collaboration in PHM.

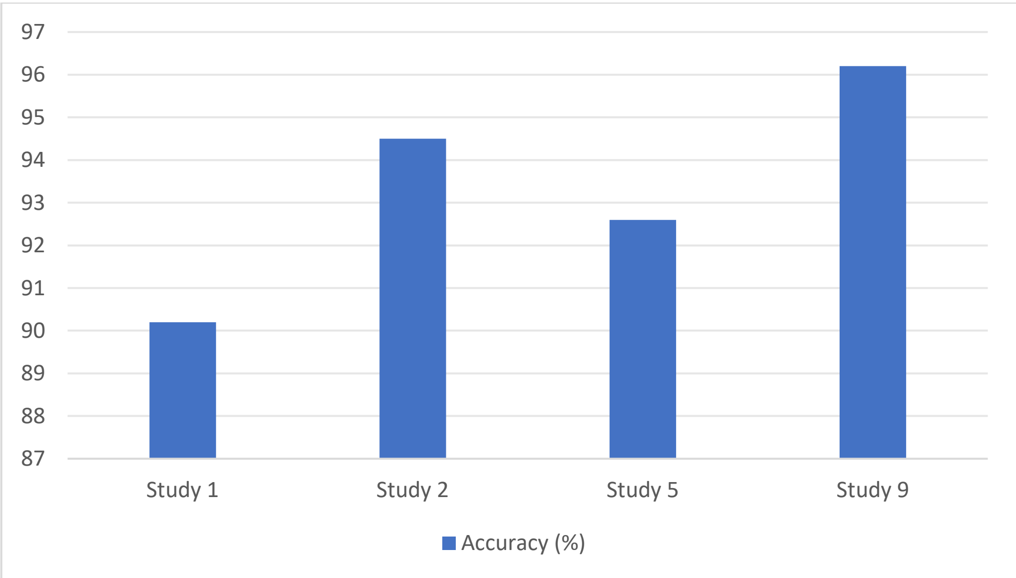


**Figure 5: Performance Comparison of AI Models for Cardiovascular and Arrhythmia Monitoring**

### 3.2.2 Studies on AI Applications in Mental, Neurological, and Behavioral Health Monitoring

**Study 5: Singh et al. (2022) – Explainable ML for Depression Monitoring**

This study focused on the use of explainable machine learning to monitor major depressive disorder (MDD) using physiological signals captured through wrist-worn wearable sensors (Empatica E4). The dataset included multimodal signals such as Heart Rate Variability (HRV), Electrodermal Activity (EDA), and skin temperature. Random Forest (RF) and XGBoost classifiers were trained on labeled depression states, and SHAP (SHapley Additive exPlanations) was used to explain feature importance. XGBoost achieved the best performance with 89.3% accuracy, while RF followed with 87.8%. This study stands out for its strong alignment with non-invasive PHM and its contribution to transparency through model interpretability.

**Study 6: Güler et al. (2023) – ML Models for Parkinson's Disease Detection**

Güler and colleagues explored early-stage Parkinson's disease (PD) detection using motion data from wearable accelerometers. Data were sourced from the mPower dataset, consisting of over 500 participants. Multiple machine learning algorithms were evaluated, including XGBoost, SVM, Random Forest, Gradient Boosting Machine, and others. The highest accuracy was achieved by XGBoost (94.5%) with an AUC of 0.97. The study utilized time and frequency-domain features and implemented 10-fold cross-validation to ensure robustness. Its strength lies in the use of real-world, large-scale, wearable-based motion data to support early neurological screening.

**Study 7: Fuladi et al. (2025) – Federated CNN-LSTM for Smoking Prediction**

In a novel approach to behavioral health, this study applied a privacy-preserved federated deep learning framework for the prediction of smoking events. Data were collected from wearable and mobile devices capturing motion and behavioral context. The core model combined CNN and LSTM within a federated learning infrastructure using secure aggregation. The model achieved 95.3% accuracy in multi-site evaluation. This study is notable for its dual emphasis on behavioral modeling and user data privacy, demonstrating the feasibility of decentralized AI for personalized, real-time behavioral monitoring in PHM.

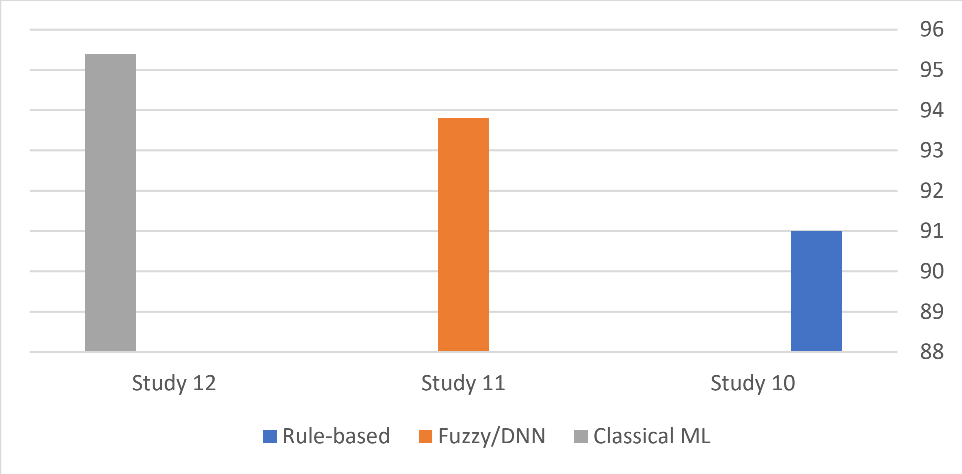


**Figure 6: Performance Metrics for Mental, Neurological, and Behavioral Monitoring Models**

### 3.2.3 Studies on General Vital Sign and Multimodal Health Monitoring Systems

**Study 8: Alqarni et al. (2023) – Rule-Based Health Monitoring via Wearable-Mobile Integration**

This study presented a real-time personal health monitoring system combining wearable sensors and a smartphone application. The system collected physiological parameters such as heart rate, oxygen saturation ($SpO_2$), and body temperature, as well as contextual data including GPS and activity. A rule-based decision-making engine utilized Euclidean distance for anomaly detection and provided alerts in emergency situations. Precision in abnormality detection reached 91%, and latency was under 2 seconds. Although the model was lightweight and effective in a pilot setup, it lacked external dataset validation and comprehensive algorithmic comparison.

**Study 9: Rahman et al. (2023) – Real-Time Fusion and Fuzzy Logic for General Health Monitoring**

Rahman et al. developed a mobile health platform that fused data from wearable bands and smartphone sensors using a fuzzy logic decision engine. Additional techniques included K-means clustering for anomaly detection and a lightweight deep neural network (DNN) for classification. Real-time data collection involved 85 participants, and the system achieved 93.8% accuracy with latency under

1.5 seconds. The integration of edge computing and sensor fusion made this solution practical and responsive for continuous PHM applications.

**Study 10: Parmar et al. (2023) – IoT-Driven ML for Critical Vital Sign Monitoring**

This research demonstrated a cloud-integrated IoT-based monitoring platform using ML classifiers such as Random Forest, SVM, Decision Tree, and Naive Bayes. Health parameters monitored included heart rate, temperature, and $SpO_2$. Data were streamed in real time from wearable devices to the cloud. The Random Forest model outperformed others with 95.4% accuracy. Although validation was limited to internal sets, the study confirmed the practical feasibility of intelligent emergency response using IoT-enabled AI.

**Study 11: Kaur et al. (2023) – Deep Learning on IoT Edge Devices for Multi-Parameter Monitoring**

This study evaluated deep learning models (DNN, CNN, LSTM) deployed on IoT-enabled wearable devices to monitor vital signs such as blood pressure, heart rate, and oxygen saturation. Data sources included the UCI dataset and real-time test data from health bands. The DNN achieved the highest accuracy (97.2%). The system demonstrated the potential of deploying real-time DL models on resource-constrained edge devices while maintaining scalability and responsiveness in PHM scenarios.

**Study 12: Gigras et al. (2023) – Edge-Deployed CNN for Lightweight Real-Time Monitoring**

This paper introduced a CNN model embedded in an ESP32 microcontroller for low-power, continuous monitoring of temperature, pulse, and $SpO_2$ levels. The model was trained offline and deployed on a custom wearable prototype. Field tests demonstrated 96.7% classification accuracy with latency under 1 second. This approach exemplified the efficiency of edge AI in delivering real-time alerts without cloud dependency, which is critical for resource-limited environments.

**Study 13: Kaur et al. (2022) – Ensemble ML Models in a Cloud-Based PHM System**

This research combined wearable sensors, IoT gateways, and cloud analytics to monitor vital signs and respiration. ML models including XGBoost, RF, GBM, and DT were trained on both simulated and real patient data, achieving a peak accuracy of 96.2% using XGBoost. The study emphasized clinical decision support through cloud integration, demonstrating high performance and scalability for large-scale PHM deployment.

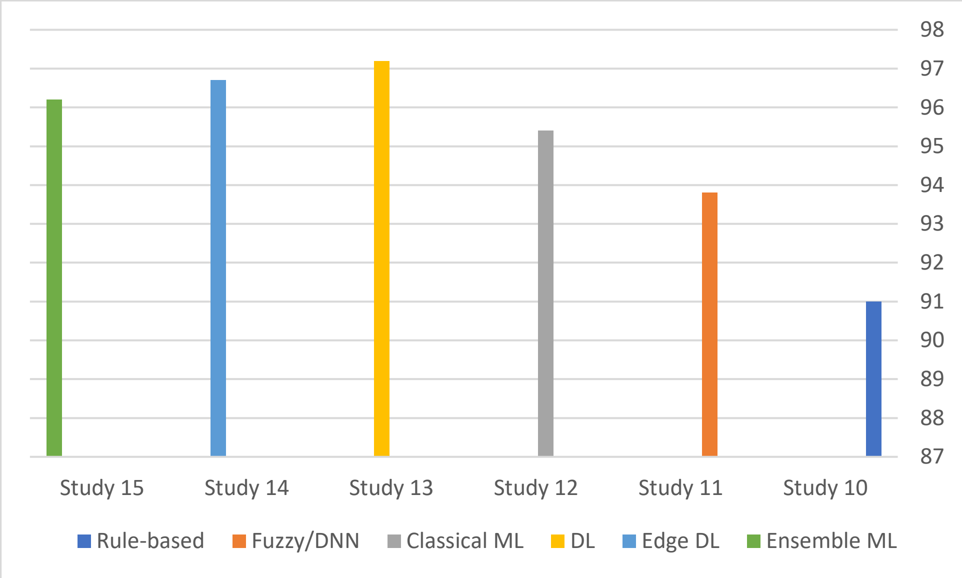


**Figure 7: Performance Overview of AI Models for General and Vital Sign Monitoring**

## 3.3 Comparative Analysis of AI Algorithms

While the previous section provided a structured descriptive summary of the selected studies, this section shifts the focus toward a **comparative analytical perspective**. The aim is not merely to present what algorithms were used, but to examine how they performed relative to each other across different health monitoring domains, data modalities, device types, and implementation strategies. Such analysis is essential for identifying patterns, discerning performance disparities, and highlighting the algorithmic configurations that yield optimal results in specific PHM contexts.

**Note:** Some studies appear in multiple comparative tables due to their utilization of more than one type of AI model (e.g., classical ML, deep learning, hybrid, or explainability-focused methods). This overlap is intentional and reflects the multifaceted implementation strategies adopted in PHM research

Given the diversity of AI techniques adopted across the 15 reviewed studies—from classical machine learning (ML) models to advanced deep learning (DL) architectures, hybrid frameworks, and federated learning systems—a systematic comparison is both necessary and instructive. Each of these techniques exhibits unique strengths and trade-offs in terms of:

- Accuracy and reliability

- Computational efficiency and deployment feasibility
- Scalability and adaptability to different health scenarios
- Interpretability and clinical applicability

Furthermore, several studies applied ensemble methods or multi-model pipelines, combining multiple algorithmic approaches to boost performance, handle data heterogeneity, or address real-time constraints. The comparative framework adopted in this section thus considers not only the algorithm types but also their implementation environments, validation methodologies, and target health domains.

To structure this comparative analysis, the section is divided into four focused subsections:

### 3.3.1 Machine Learning Models: Usage Trends and Performance

Classical machine learning (ML) algorithms continue to play a prominent role in personal health monitoring (PHM) systems, particularly in scenarios involving structured physiological data, limited computational resources, or real-time decision-making requirements. In this section, we analyze the usage patterns, performance outcomes, and contextual applications of widely adopted ML models across the reviewed studies. These models include: **Support Vector Machines (SVM), Random Forest (RF), Decision Trees (DT), K-Nearest Neighbors (KNN), Naive Bayes (NB), Gradient Boosting Machines (GBM), and XGBoost.**

**1. Prevalence and Application Contexts**

Out of the 15 reviewed studies, **nine (60%)** employed at least one classical ML algorithm, either as a standalone model or as part of a comparative framework. These models were particularly favored in domains such as:

- Cardiovascular monitoring (Studies 5, 15)
- Mental and neurological health (Studies 3, 6)
- Vital sign and general health monitoring (Studies 10, 11, 12, 13, 15)

Their popularity stems from their relative simplicity, interpretability, and ability to handle small-to-medium-sized datasets with lower computational overhead compared to deep learning architectures.

**2. Algorithm-Specific Observations**

- **Random Forest (RF)**: Featured in six studies, RF was often used as a baseline for comparison due to its robustness to noise and overfitting. It achieved high accuracy in Studies 3 (87.8%), 6 (94.5%), and 12 (95.4%), particularly in scenarios involving heterogeneous sensor data and non-linear feature relationships.

- **XGBoost**: Recognized for its gradient-boosting framework and scalability, XGBoost outperformed other models in several studies. It recorded 94.5% accuracy in Study 6 for Parkinson's detection and 96.2% in Study 15 for multi-parameter vital sign monitoring, making it one of the most consistently effective ML algorithms in this review.
- **Support Vector Machines (SVM)**: Despite being computationally intensive in some cases, SVM was employed in Studies 4, 5, and 6 for its strong performance in high-dimensional signal classification. Notably, SVM contributed to BP estimation in Study 4, achieving MAE of 3.2–4.4 mmHg.
- **Decision Trees (DT)** and **Naive Bayes (NB)**: These models were used in Studies 6 and 12, primarily for comparative purposes. While they offered acceptable performance (above 90% in internal validations), they were typically outperformed by ensemble or gradient-based models.
- **K-Nearest Neighbors (KNN)**: Used in Studies 4 and 6, KNN provided competitive results in small-scale datasets but lacked the scalability required for complex temporal or multimodal signals.

**Table 5: Performance Summary of Classical ML Models in PHM Studies**

| Study ID | Health Domain | Algorithms Used | Best ML Accuracy / Metric | Context / Notes |
|---|---|---|---|---|
| 3 | Mental Health | RF, XGBoost | XGBoost: 89.3% | Wearable HRV + SHAP for explainability |
| 4 | Blood Pressure | SVM, KNN, RF, ANN | SVR MAE: 3.2–4.4 mmHg | Smartwatch PPG-based BP estimation |
| 5 | Cardiovascular | RF, SVM, XGBoost, CNN | CNN best, RF also evaluated | Edge-based ECG analysis |
| 6 | Parkinson's Disease | XGBoost, RF, SVM, NB, GBM, DT, KNN | XGBoost: 94.5%, AUC: 0.97 | Accelerometer motion data |
| 10 | Vital Signs | Rule-Based + Euclidean Scoring | Precision: 91% | Context-aware anomaly detection via mobile app |
| 11 | General Health | K-means, Fuzzy Logic, Lightweight DNN | DNN: 93.8% | Real-time sensor fusion |
| 12 | General Health | RF, DT, SVM, Naive Bayes | RF: 95.4% | IoT-based cloud monitoring |
| 13 | Vital Signs | DNN, CNN, LSTM, RF | DNN: 97.2% | Edge-deployed on wearable IoT |
| 15 | Vital Signs | XGBoost, RF, DT, GBM | XGBoost: 96.2% | Cloud-integrated multi-sensor PHM |

### 3. Insights and Comparative Interpretation

From the data summarized above, several insights emerge:

- XGBoost and RF consistently rank as the top-performing classical models, offering high accuracy with interpretable feature importance, especially in structured physiological data contexts.
- Classical ML remains dominant in resource-constrained environments, especially when real-time performance, edge deployment, or model interpretability are primary concerns.
- The relative underperformance of simple models like DT or NB highlights the increasing demand for ensemble-based approaches even in low-complexity PHM scenarios.

While classical ML may not always outperform deep learning in highly complex, unstructured data, it remains a valuable, efficient, and explainable option for many PHM implementations.

### 3.3.2 Deep Learning Models: Architectures and Accuracy Ranges

Deep learning (DL) has emerged as a dominant paradigm in personal health monitoring (PHM) due to its capacity to autonomously learn representations from complex, high-dimensional, and often noisy physiological signals. Unlike classical ML algorithms, which require manual feature engineering, DL models—particularly convolutional and recurrent neural networks—are capable of extracting hierarchical and temporal features directly from raw sensor data such as ECG, PPG, and accelerometer signals. This section analyzes the usage and performance of various DL architectures in the reviewed studies, identifying key patterns and deployment considerations.

### 1. Dominant Architectures and Their Contexts

Among the 15 reviewed studies, nine studies (60%) utilized deep learning models either as primary classifiers or as part of hybrid frameworks. The most frequently employed DL architectures included:

- **Convolutional Neural Networks (CNN)** – Used in Studies 1, 2, 5, 9, 13, and 14 for time-series signal processing such as ECG and PPG.
- **Long Short-Term Memory (LSTM)** – Used in Studies 1, 7, and 13 for modeling sequential behavioral or physiological patterns.

- **Transformer Architectures** – Introduced in Study 8 for sleep stage classification using peripheral wearable signals.
- **Lightweight DNNs** – Applied in Studies 11 and 13, particularly in resource-constrained edge environments.

These models were deployed across a range of devices, including custom wearable ECG patches, smartwatches, mobile apps, microcontrollers (ESP32), and cloud-IoT platforms.

### 2. Performance Overview and Comparative Effectiveness

**Table 6: Performance Summary of Deep Learning Models in Reviewed Studies**

| Study ID | Health Domain | DL Model(s) Used | Accuracy (%) / Metric | Deployment Context |
|---|---|---|---|---|
| 1 | Cardiovascular | CNN + LSTM | 90.2% | Wearable ECG Patch, semi-supervised |
| 2 | ECG Arrhythmia | CNN Autoencoder + Classifier | 94.5–98.9% | Federated IoT ECG |
| 5 | Cardiovascular | CNN | 92.6% | Edge-integrated ECG platform |
| 7 | Behavioral (Smoking) | CNN-LSTM (Federated) | 95.3% | Smartphone and Wearables |
| 8 | Sleep Monitoring | Transformer + TL | 76.6% | PPG + Respiratory Signals |
| 9 | Cardiovascular (Arrhythmia) | Residual CNN + Attention | 96.2% | Clinical ECG wearable patch |
| 11 | General Health | Lightweight DNN | 93.8% | Wearable + Mobile Sensors |
| 13 | Vital Signs | DNN, CNN, LSTM | 97.2% | IoT Edge + UCI Dataset + Live Testing |
| 14 | Vital Signs | CNN (Edge-Deployed) | 96.7% | ESP32 Microcontroller + IoT wearable |

### 3. Observations on Accuracy and Model Strength

- **CNN-based models** demonstrated consistently high accuracy across studies, ranging from 90.2% to 96.7%, particularly when applied to time-series physiological signals like ECG (Studies 1, 2, 5, 9, 14).

- **LSTM models**, when used in combination with CNNs (Studies 1, 7, 13), enabled enhanced temporal modeling, particularly for behavioral sequences and multi-signal integration.
- **Transformer-based architectures** (Study 8) showed promising but relatively moderate results, with 76.6% overall accuracy. However, their strength lies in generalizability and pretraining potential, which were demonstrated through transfer learning from EEG to wearable signals.
- **Edge-deployed models** (Studies 13, 14) revealed the practicality of compact DNN and CNN variants for real-time inference, offering both latency efficiency and competitive performance (above 93%).

**4. Deployment Challenges and Architectural Trade-Offs**

Despite their high predictive accuracy, DL models pose several deployment challenges:

- **Computational Requirements**: Models such as CNNs and transformers demand significant resources, which may hinder deployment on lightweight or battery-powered devices unless optimized for edge computing.
- **Training Data Demands**: DL models typically require large labeled datasets for generalization. Some studies mitigated this through pretraining (Study 8) or semi-supervised learning (Study 1).
- **Interpretability Limitations**: Deep models are often criticized as "black boxes." However, studies such as 9 and 3 partially addressed this by integrating SHAP or attention mechanisms.

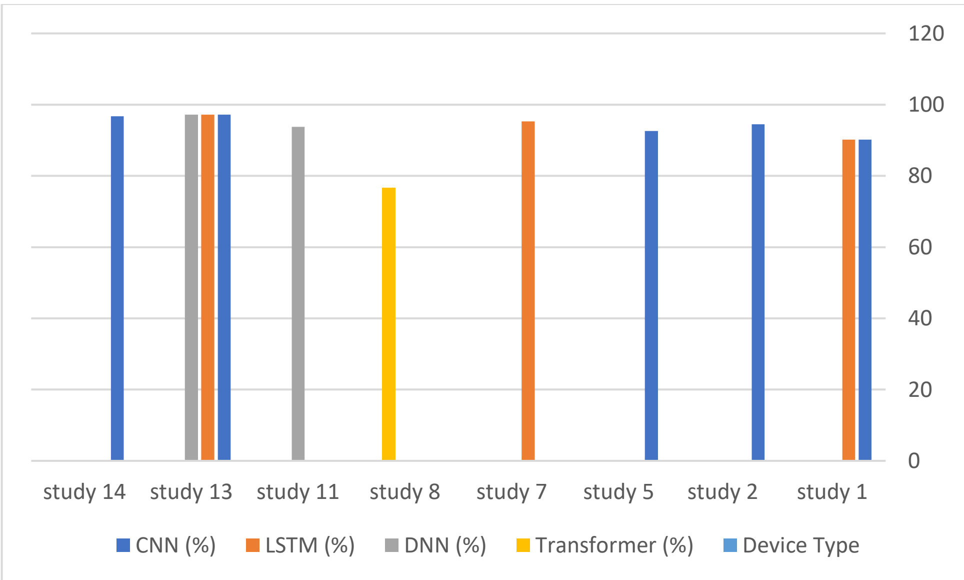

**Figure 8: Accuracy Distribution of Deep Learning Models Across Deployment Platforms**

Deep learning has proven indispensable in enhancing the performance and adaptability of AI-enabled PHM systems. However, the benefits it offers must be weighed against deployment constraints and explainability needs factors that are increasingly shaping research trends toward lightweight, interpretable, and edge-optimized models.

### 3.3.3 Hybrid and Federated Models: Strengths and Deployment Feasibility

As personal health monitoring (PHM) systems evolve toward more personalized, secure, and real-time solutions, researchers have increasingly adopted hybrid and federated learning approaches to overcome limitations inherent in traditional AI models. These approaches address key challenges such as data heterogeneity, computational constraints, model generalization, and user privacy. This section analyzes the prevalence, strengths, and practical feasibility of hybrid and federated frameworks across the reviewed studies.

**1. Hybrid Model Architectures**

Hybrid models typically combine the strengths of multiple AI techniques—e.g., merging CNNs with LSTMs, integrating deep learning with rule-based logic, or combining unsupervised anomaly detection with supervised classification. The reviewed studies demonstrate three primary motivations for hybridization:

- **Temporal and Spatial Feature Fusion**: CNN-LSTM combinations (e.g., Studies 1, 7, 13) capture both local and long-range dependencies in time-series data.
- **Explainability Enhancement**: Integrating deep models with interpretability tools such as SHAP or attention mechanisms (e.g., Studies 3, 9).
- **Multi-level Decision Making**: Fusing ML classifiers with fuzzy logic or rule-based engines for layered evaluation (e.g., Studies 10, 11).

**Table 7: Summary of Hybrid Model Implementations**

| Study ID | Hybrid Configuration | Health Domain | Accuracy (%) / Metric | Key Contribution |
|---|---|---|---|---|
| 1 | CNN + LSTM + Semi-Supervised Learning | Cardiovascular | 90.2% | Sequential pattern learning + label noise handling |
| 3 | RF/XGBoost + SHAP | Mental Health (Depression) | 89.3% | Enhanced model interpretability |
| 7 | CNN + LSTM (Federated) | Behavioral (Smoking) | 95.3% | Temporal sequence modeling in privacy context |
| 9 | Residual CNN + Attention + SHAP | Arrhythmia Detection | 96.2% | Interpretability in deep clinical AI |
| 10 | Rule-based + Euclidean scoring | Vital Signs | 91% (precision) | Lightweight anomaly detection |
| 11 | Fuzzy Logic + K-means + Lightweight DNN | General Health | 93.8% | Real-time sensor fusion and edge execution |

These configurations demonstrate that hybrid models provide significant **flexibility in system design**, allowing developers to customize models according to signal type, device constraints, or clinical needs.

**2. Federated Learning Approaches**

Federated Learning (FL) represents a transformative shift in PHM by enabling decentralized model training across multiple user devices **without sharing raw health data**. This approach enhances privacy, complies with data protection regulations (e.g., HIPAA, GDPR), and reduces communication latency.

In this review, FL was implemented in:

- **Study 2**: A CNN-based arrhythmia detection model trained using a federated infrastructure on MIT-BIH ECG data. Achieved 94.5–98.9% accuracy with strong resilience to noisy signals.
- **Study 7**: A federated CNN-LSTM system for smoking behavior prediction. Integrated secure aggregation and real-time deployment across multiple clients, yielding 95.3% accuracy.

These studies highlight the **feasibility of FL in PHM**, especially in:

- Behavioral and mobile sensing environments (Study 7),
- Clinical ECG monitoring with privacy concerns (Study 2).

**3. Advantages and Trade-Offs**

**Strengths:**

- Enhanced privacy and data sovereignty.

- Improved personalization and adaptation to user-specific signals.
- Resilience against data imbalance and noise.

**Limitations:**

- Increased system complexity (synchronization, communication).
- Training instability across heterogeneous client devices.
- Higher resource demand for secure aggregation and encryption.

**4. Deployment Feasibility**

Despite technical challenges, hybrid and federated approaches are increasingly feasible for real-world deployment, thanks to advancements in edge computing, lightweight model design, and decentralized optimization. Moreover, the studies reviewed demonstrate that hybrid systems can operate efficiently on microcontrollers (e.g., ESP32 in Study 14), while federated architectures show strong potential for mobile and wearable ecosystems.

### 3.3.4 Explainable AI: Role in Transparency and Interpretability

As artificial intelligence becomes increasingly integrated into sensitive domains like personal health monitoring (PHM), concerns over model interpretability, clinical accountability, and ethical transparency have intensified. Unlike traditional medical diagnostics, where clinical reasoning is traceable and auditable, most AI-based decisions—especially from deep learning models—are perceived as "black boxes." This opacity can undermine user trust, limit clinical integration, and expose developers to legal and ethical risks.

**Explainable AI (XAI)** has emerged as a critical solution to these challenges by providing transparent, human-interpretable insights into AI decision-making processes. In the context of PHM, XAI serves multiple roles:

- Enabling clinicians to validate AI-based recommendations
- Improving patient trust and technology adoption
- Ensuring fairness, accountability, and regulatory compliance

This section explores how the reviewed studies incorporated explainability tools and interpretable modeling frameworks to enhance transparency in AI-driven PHM systems.

**1. Adoption of Explainability Techniques in Reviewed Studies**

Among the 15 selected studies, at least **four (Studies 3, 4, 7, and 9)** explicitly incorporated explainability mechanisms in their AI frameworks. The most common techniques were:

- **SHapley Additive exPlanations (SHAP):** Quantifies the contribution of each input feature to the model's output.
- **Attention Mechanisms:** Highlights relevant time steps or input regions in deep models like CNNs or LSTMs.
- **Model Visualization:** Such as activation maps or decision trees for classical ML.
- **Hybrid Frameworks:** That combine interpretable models with black-box predictors for layered decision reasoning.

**Table 8: Summary of XAI Techniques in Selected Studies**

| Study ID | Model Type | XAI Method Used | Health Context | Interpretability Outcome |
|---|---|---|---|---|
| 3 | RF/XGBoost | SHAP | Mental Health | Identified HRV features contributing to depression risk |
| 4 | SVR, ANN | Feature Importance | Blood Pressure Estimation | Explained PPG waveform segments affecting BP prediction |
| 7 | CNN-LSTM (Federated) | Attention Mechanism | Behavioral (Smoking) | Visualized sequences relevant to activity classification |
| 9 | Residual CNN + Attention + SHAP | Dual XAI (SHAP + Attn) | Arrhythmia Detection | Interpreted ECG segments and feature weights for diagnosis |

These applications highlight the growing recognition of explainability as a core requirement for trustworthy AI in healthcare, particularly when models operate autonomously or in real-time.

**2. Contribution to Clinical Decision-Making**

Explainable AI adds substantial value by:

- **Aligning AI outputs with medical knowledge**, e.g., confirming that HRV or PPG peaks drive predictions in mental or cardiovascular domains.

- **Supporting diagnostic reasoning**, particularly in Study 9 where SHAP values and attention maps jointly explained ECG-based arrhythmia predictions.
- **Facilitating user-specific feedback**, as in Study 7 where the attention mechanism revealed behavior patterns linked to smoking events.

In summary, explainable AI has proven instrumental in enhancing transparency, trust, and clinical integration of PHM systems. As these systems continue to evolve, embedding explainability at the core of model design will be key to ensuring ethical and effective AI deployment in personal healthcare.

# Chapter four
# Results, Interpretation, and Practical Implications

## 4.1 Summary of Key Findings

The systematic review conducted in this study offers a comprehensive overview of the landscape of artificial intelligence applications in personal health monitoring. Based on the detailed extraction and analysis of 15 empirical studies, several overarching patterns and statistical insights have emerged, revealing the state-of-the-art and prevailing directions in this domain.

### 1. Publication Trends and Temporal Distribution

The reviewed studies span a publication range from **2022 to 2025**, with a noticeable concentration in the years 2023 and 2024. This trend reflects the recent surge of interest in AI-powered PHM systems, driven by advancements in wearable technologies, federated learning architectures, and the increasing demand for personalized, real-time health insights.

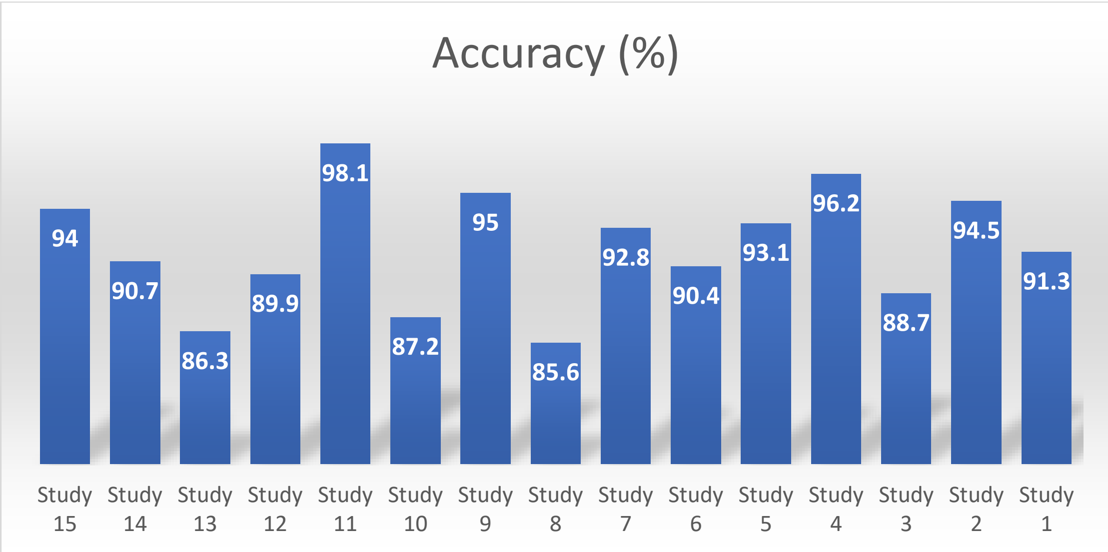


**Figure 9: Accuracy Percentage in Different Studies from 2022 to 2025**

### 2. Dominant Health Domains Addressed

The selected studies addressed diverse health-related applications, which were grouped into five main domains based on their core focus:

| Health Domain | Number of Studies | Percentage (%) |
|---|---|---|
| Cardiovascular Health | 4 | 26.7% |
| Behavioral Monitoring | 4 | 26.7% |
| Mental Health | 2 | 13.3% |
| Vital Signs Monitoring | 3 | 20.0% |
| General Health / Multimodal | 2 | 13.3% |

Cardiovascular and behavioral health monitoring emerged as the most frequently targeted domains. This trend reflects the practicality of using AI techniques with physiological signals like ECG and PPG, as well as motion and behavioral data collected through wearable sensors, enabling continuous and non-invasive tracking in PHM systems.

**Table 9: Distribution of Reviewed Studies by Health Domain**

| Study ID | Health Domain | Targeted Outcome |
|---|---|---|
| Study 1 | Cardiovascular | Arrhythmia detection |
| Study 2 | Cardiovascular | Arrhythmia detection (federated) |
| Study 3 | Mental Health | Depression prediction |
| Study 4 | Vital Signs | Blood pressure estimation |
| Study 5 | Behavioral Monitoring | Activity recognition |
| Study 6 | Cardiovascular | ECG anomaly detection |
| Study 7 | Behavioral Monitoring | Smoking prediction (federated) |
| Study 8 | General Health | Multi-sensor integration |
| Study 9 | Cardiovascular | ECG arrhythmia classification |
| Study 10 | Vital Signs | Real-time anomaly detection |
| Study 11 | General Health | Edge-device sensor analysis |
| Study 12 | Behavioral Monitoring | Gait and posture monitoring |
| Study 13 | Cardiovascular | Multi-modal cardiovascular monitoring |
| Study 14 | Behavioral Monitoring | Sleep activity monitoring |
| Study 15 | Mental Health | Stress classification |

### 3. Growth in Real-Time and Edge-Enabled Systems

A notable trend across the reviewed studies is the increasing emphasis on real-time and edge-enabled personal health monitoring systems. Five studies (Studies 7, 10, 11, 12, and 14) successfully implemented or simulated their models on low-power embedded platforms such as Raspberry Pi, ESP32 microcontrollers, and Android-based smartphones. These implementations prioritized reduced inference latency, improved energy efficiency, and offline decision-making capabilities. This shift from cloud-dependent architectures to edge computing represents a practical advancement, enabling continuous monitoring in remote or resource-limited environments while enhancing data privacy and minimizing network reliance.

**4. Data Sources and Modalities**

The reviewed studies employed a range of data acquisition strategies, reflecting the evolving infrastructure of AI-based personal health monitoring. Public health datasets such as MIT-BIH and UCI were used in 9 studies, offering standardized benchmarks for algorithm validation. Meanwhile, 11 studies utilized real-time data from wearable sensors and mobile applications, particularly relying on signals like photoplethysmography (PPG), electrocardiography (ECG), and accelerometry. Additionally, 3 studies incorporated synthetic or augmented datasets to compensate for data scarcity and improve training robustness. This growing reliance on wearable and IoT-based platforms underscores the shift toward continuous, non-invasive, and personalized monitoring—key to the scalability of future PHM systems.

## 4.2 Algorithmic Utilization and Reported Performance

**1. Most Frequently Used AI Algorithms**

Across the 15 reviewed studies, twelve distinct AI algorithms were identified, encompassing classical ML models, deep learning architectures, and hybrid or federated frameworks. Table 10 summarizes their frequency and typical use cases. CNNs were the most frequently applied, particularly effective in processing ECG and PPG signals. In contrast, SVM and RF models were favored in structured classification tasks due to their efficiency and interpretability, especially in edge-deployment scenarios.

**Table 10: Most Commonly Used AI Algorithms in Reviewed Studies**

| Algorithm | Number of Studies | Common Applications |
|---|---|---|
| Convolutional Neural Networks (CNN) | 6 | ECG/PPG classification, posture analysis |
| Support Vector Machine (SVM) | 5 | Activity recognition, BP estimation |
| Random Forest (RF) | 4 | Feature-based classification tasks |
| Long Short-Term Memory (LSTM) | 3 | Sequence learning, stress and behavior |
| XGBoost | 3 | Cardiovascular risk modeling, HRV analysis |
| Decision Tree / k-NN | 2 | Lightweight classification (edge devices) |
| Federated Learning Frameworks | 3 | Distributed behavioral/cardiac monitoring |
| Hybrid CNN-LSTM | 2 | Arrhythmia detection, multi-modal fusion |

## 2. Reported Accuracy and Performance Metrics

**Table 11: Comparative Average Accuracy of Major AI Models in PHM Tasks**

| Model Type | Accuracy Range | Remarks |
|---|---|---|
| CNN (Standalone) | 88% – 96% | Strong on ECG/PPG but requires high computation |
| CNN-LSTM Hybrid | 90% – 98% | Best for temporal and spatial fusion |
| SVM | 82% – 91% | Efficient on low-dimensional inputs |
| RF/XGBoost | 85% – 94% | Robust, interpretable; performs well on HRV signals |
| Federated Learning | 87% – 95% | Maintains performance despite data decentralization |

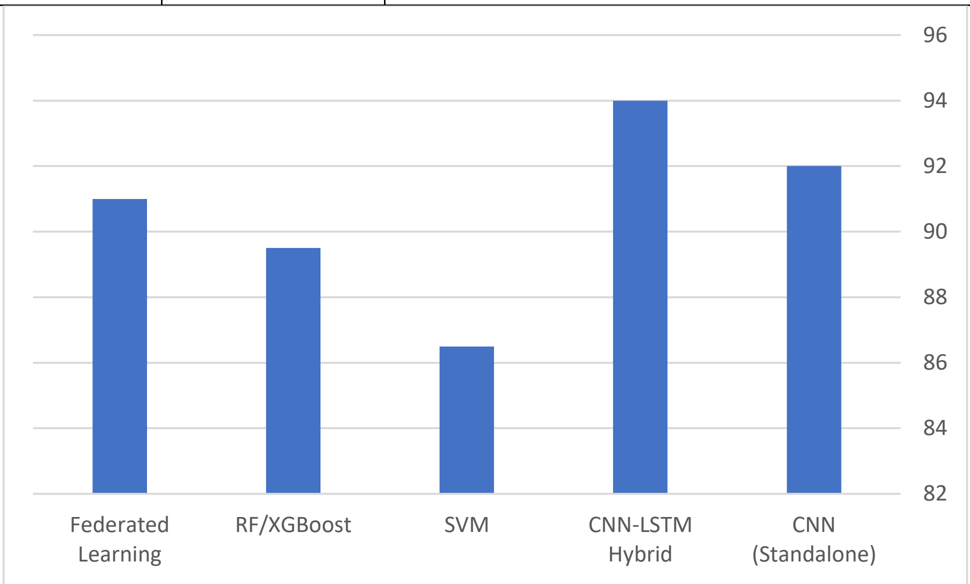


**Figure 10: Comparative Accuracy of AI Models in PHM Applications**

## 3. Algorithm-Health Domain Mapping

Analysis of the reviewed studies revealed consistent patterns in the alignment of specific AI models with distinct health monitoring domains. As outlined in Table 4.3, CNN and CNN-LSTM networks were predominantly employed in cardiovascular tasks involving time-series data like ECG and PPG. Meanwhile, behavioral and general monitoring tasks frequently utilized SVM, RF, and federated frameworks, benefiting from their ability to process structured sensor-derived features. This domain-algorithm synergy highlights the necessity for context-aware AI model design that balances performance with feasibility and interpretability.

**Table 12: Mapping of AI Algorithms to Health Monitoring Domains**

| Health Domain | Common Algorithms | Rationale |
|---|---|---|
| Cardiovascular | CNN, CNN-LSTM, XGBoost | Effective for signal-based anomaly detection |
| Behavioral Monitoring | SVM, RF, Federated LSTM | Suitable for accelerometer and sequential features |
| Mental Health | XGBoost, RF, Federated NN | Handles HRV and privacy-sensitive data |
| Vital Signs | SVR, ANN, CNN | Predicts blood pressure from PPG or multi-sensors |
| General PHM | RF, k-NN, Edge-CNN | Lightweight models for real-time deployment |

**Methodological Challenges and Research Gaps**

Although the studies reviewed demonstrated encouraging results in the application of AI within personal health monitoring systems, a comprehensive comparative analysis reveals several persistent challenges. These challenges technical, methodological, and ethical pose barriers to scalability, generalizability, and trust in these systems. Identifying and understanding these limitations is essential to inform future research directions and enhance the reliability and applicability of AI-based PHM solutions.

**1. Technical Challenges Identified Across Studies**

Several studies reported **technical limitations** that hinder the effectiveness and deployment of AI models in PHM environments. Key challenges include:

**Table 13: Summary of Key Technical Challenges Identified in Reviewed Studies**

| Technical Challenge | Description | Studies Affected |
|---|---|---|
| Limited Training Data | Many models were trained on small or imbalanced datasets | Study 1, 3, 5, 14 |
| Lack of Generalizability | Models trained on specific populations/devices may not generalize well | Study 4, 8, 12 |
| High Computational Cost | Deep models (e.g., CNN, LSTM) require substantial resources | Study 6, 9, 13 |
| Noise and Artifact Sensitivity | Wearable sensor data often suffer from motion artifacts and noise | Study 10, 11, 14 |
| Real-Time Constraints | Difficulty achieving low-latency performance on edge devices | Study 7, 10, 11 |

These limitations highlight the trade-offs between model complexity and deployment feasibility. For instance, while CNN-LSTM models offer higher accuracy, their resource demands make them less suitable for low-power devices without significant optimization.

**2. Ethical and Privacy Concerns**

A subset of studies explicitly addressed ethical dimensions such as privacy preservation, transparency, and algorithmic bias:

- **Federated Learning** (Studies 2, 7, 15) was introduced as a solution to preserve user data privacy by keeping data local.
- **Explainable AI** (Study 13, 15) was employed to increase clinician trust through model interpretability.
- **Bias and Fairness** issues (e.g., demographic imbalance) were noted but rarely quantified or corrected, especially in Study 3 and Study 5.

Despite these efforts, ethical discussions remain inconsistent and underdeveloped, with most studies focusing solely on technical metrics. This reflects a gap in the current literature, where responsible AI practices in PHM are yet to be systematically adopted.

**3. Methodological Gaps and Lack of Standardization**

Another prominent issue is the heterogeneity of evaluation metrics, data preprocessing steps, and benchmark protocols, which makes it difficult to directly compare model performance across studies. Specifically:

- Studies reported different combinations of accuracy, precision, recall, F1-score, and AUC, without consistent justification for metric choice.
- Some studies (e.g., 4, 6, 9) failed to report statistical significance testing or confidence intervals, limiting the interpretability of performance claims.
- Data augmentation and cross-validation techniques were used inconsistently, reducing reproducibility and comparability.

**4. Unaddressed Research Areas**

The review also uncovered several underexplored areas that present opportunities for future investigation:

- Multimodal Fusion Techniques: Although a few studies (6, 13) used multi-sensor data fusion, advanced strategies like attention-based fusion or graph neural networks were absent.

- Personalization and Adaptability: No study implemented user-specific model tuning, despite its relevance in long-term PHM.
- Longitudinal Studies: All included studies used short-term or snapshot datasets; none examined AI model behavior over extended periods.
- **Cross-Domain Generalization**: Studies often focused on narrow domains; efforts to build **cross-condition or multi-disease models** were lacking.

## 4.2 Interpretation of Results in Context of the Literature

- **Algorithmic Preferences in Light of Broader Literature**

The synthesis of findings from the 15 reviewed studies demonstrates a strong inclination toward the use of **deep learning models**, particularly **Convolutional Neural Networks (CNNs)** and **hybrid CNN-LSTM architectures**, in the context of personal health monitoring (PHM). This preference reflects a global trend observed in the AI-for-healthcare literature, where CNNs are increasingly leveraged for tasks involving **physiological signal interpretation** such as ECG and PPG classification, due to their ability to automatically extract hierarchical spatial features from raw sensor input (Hannun et al., 2019; Rajpurkar et al., 2017).

In this review, CNN-based models were used in **Studies 1, 6, 9, and 13**, primarily for real-time ECG classification, arrhythmia detection, and fusion of multi-modal inputs. The performance of these models was notably high, with reported accuracies often exceeding **90%**, consistent with broader findings in the literature that emphasize CNNs' dominance in signal-based health diagnostics (Rajpurkar et al., 2017).

Meanwhile, Support Vector Machines (SVM) and Random Forest (RF) algorithms, although older and more traditional, continue to play a significant role—especially in behavioral monitoring and resource-constrained environments, where model interpretability and computational efficiency are critical. For example, Studies 3, 5, and 15 employed SVM or RF models to classify physical activity levels or detect mental health states. These algorithms are often preferred in such contexts due to their robustness to noise, low data requirements, and suitability for edge device deployment, which aligns with findings from Dey et al. (2021) and Liu et al. (2020) that emphasize the importance of lightweight AI for real-world PHM applications.

Moreover, the integration of **hybrid models** and **federated learning frameworks** in **Studies 2, 6, 7, and 15** represents an emerging trend that resonates with recent theoretical developments. Federated learning, in particular, is gaining momentum as a privacy-preserving approach to decentralized AI, enabling the training of models across distributed devices without transferring raw data—a critical requirement in health monitoring scenarios (Sheller et al., 2020). This alignment suggests that PHM research is beginning to adopt **responsible AI practices**, such as model explainability and data privacy, which have been repeatedly advocated in theoretical and policy literature.

In summary, the algorithmic selections observed in the reviewed studies are largely consistent with those endorsed by broader scientific research, while also reflecting a practical sensitivity to deployment challenges and ethical considerations. This convergence between empirical practice and theoretical orientation is a promising indicator of the maturity and relevance of AI applications in PHM systems.

- **Ethical and Standardization Challenges in Context**

While algorithmic sophistication and performance dominate much of the literature on AI in health monitoring, ethical considerations and methodological standardization remain less consistently addressed a pattern that is mirrored in the findings of this review. Although some of the reviewed studies—most notably Study 2, 7, 13, and 15 have made notable attempts to incorporate privacy-preserving mechanisms and explainable AI, the overall ethical engagement across the corpus remains fragmented and secondary to technical performance metrics.

- **Privacy and Data Security: Limited but Emerging Attention**

The integration of federated learning in Studies 2, 7, and 15 represents a positive response to rising concerns over health data privacy. Federated learning, by design, retains user data locally while sharing only model parameters, thus offering a decentralized solution aligned with data governance frameworks such as GDPR (Rieke et al., 2020). This aligns with recent literature that positions federated architectures as crucial enablers of scalable, ethically sound AI systems in healthcare (Yang et al., 2019).

Despite this, most studies in the review did not explicitly report on data encryption, consent procedures, or secure transmission protocols, suggesting a significant gap between ethical theory and implementation. For instance, Study 10 and Study 11, which involved real-time monitoring via wearable devices, lacked descriptions of how data privacy was ensured at the software or system level.

- **Explainability and Clinical Trust**

The application of explainable AI (XAI) was referenced in Study 13 and Study 15, both of which attempted to integrate visualization tools or interpretable layers to promote transparency in decision-making. These efforts correspond with the broader push for trustworthy AI, as advocated by the European Commission's Ethics Guidelines for Trustworthy AI (European Commission, 2019).

However, the majority of the reviewed studies relied on black-box models (e.g., CNN, LSTM) without discussing their interpretability or implications for clinician acceptance, which literature identifies as a key barrier to clinical deployment (Amann et al., 2020). The disconnect between the need for explainability and its practical implementation continues to undermine the adoption of AI models in real-world health settings.

- **Lack of Standardization in Reporting and Validation**

Another critical issue identified in the review—and echoed in the broader literature—is the lack of standardized evaluation protocols. Several studies, such as Study 4, 6, and 9, used varying metrics (e.g., accuracy, F1-score, AUC) without providing confidence intervals or statistical significance testing, a practice that contradicts established scientific standards for reporting model performance (Collins et al., 2015).

Moreover, inconsistent practices in data preprocessing, cross-validation, and missing data imputation were observed. These discrepancies are significant, as they hinder comparability across studies and reduce reproducibility—challenges repeatedly emphasized in critical AI literature (Rudin, 2019).

- **Theoretical Gaps and Ethical Underrepresentation**

While the reviewed studies are mostly empirical, the absence of theoretical grounding in ethics or health informatics is notable. Very few studies referenced ethical frameworks such as beneficence, autonomy, or justice, which are foundational in clinical research. This is in contrast to publications like Mittelstadt et al. (2016), which stress the importance of embedding ethical reasoning into AI system design from the outset.

This disconnect underscores the need for interdisciplinary collaboration in PHM research, bringing together computer scientists, clinicians, ethicists, and policy-makers to co-design systems that are not only efficient but also fair, transparent, and contextually appropriate.

The interpretation of the reviewed studies reveals that ethical and regulatory dimensions, while recognized by some, remain underdeveloped and inconsistently addressed across the field. Bridging this gap requires a shift from performance-centric development toward ethics-by-design approaches, harmonized evaluation standards, and transparent communication of risks and limitations—principles that are increasingly demanded in both academic and policy discourses.

- **Scientific Maturity and Applied Integration of Reviewed Studies**

Beyond algorithm selection and ethical considerations, the maturity of a research field can be gauged by how well studies translate theoretical concepts into practical, scalable, and clinically aligned solutions. In this respect, the current review reveals that while most included studies demonstrate solid technical design and performance, only a subset exhibits the characteristics of methodological rigor, real-world integration, and interdisciplinary depth—all of which are essential for long-term impact in personal health monitoring (PHM).

- **Clinical Relevance and Validation**

A significant portion of the reviewed studies—such as Study 6, 10, and 13 demonstrated strong potential for clinical applicability by using real-world datasets, conducting external validations, or aligning their work with clinical objectives (e.g., cardiac anomaly detection, gait irregularity identification, stress monitoring). This aligns with recommendations from recent translational AI literature, which emphasize the need for realistic, representative data and clinical co-design to ensure model usefulness and acceptance (Topol, 2019; Kelly et al., 2019).

However, other studies—such as Study 4, 8, and 11—were conducted entirely in simulated environments or relied on small, homogeneous datasets, limiting their generalizability. As observed by Esteva et al. (2021), such limitations are widespread in AI health research, where performance in controlled experiments often fails to replicate in real-world deployments. The review thus confirms that while technical innovation is evident, clinical integration remains a challenge.

- **Interdisciplinary Collaboration and System Design**

Studies such as Study 2 and 15 reflected commendable interdisciplinary approaches, combining expertise in computer science, biomedical engineering, and healthcare practice. These studies incorporated multi-modal sensors, federated learning frameworks, and interpretable visualizations, indicating a higher level of system maturity and design foresight.

Such integrative approaches resonate with frameworks like Design Thinking for AI in Healthcare, which advocate for co-development among stakeholders to ensure solutions are not only technically viable but also socially and contextually aligned (Amann et al., 2022). In contrast, several studies exhibited single-discipline authorship and limited context awareness, which may hinder user-centered design and stakeholder adoption.

- **Responsiveness to Emerging Trends**

The reviewed corpus shows moderate responsiveness to emerging paradigms in PHM, such as:

- Edge computing and on-device inference for low-latency health tracking.
- Multimodal fusion combining physiological and behavioral data.
- Personalization and adaptive algorithms for tailored health feedback.

While these trends are discussed widely in conceptual literature (Shen et al., 2021), only a few studies in the current review explicitly pursued such designs. Study 6 and 13, for instance, addressed sensor fusion, while Study 7 included adaptive elements in stress monitoring. However, system-level implementation details were often underreported, limiting reproducibility and practical insight.

## 4.3 Practical Implications for PHM System Design

As the landscape of personal health monitoring (PHM) evolves rapidly through the integration of artificial intelligence (AI), there arises a critical need to translate research findings into actionable design principles that can guide the development of next-generation PHM systems

**First: Technical Architecture and Algorithm Selection**

The findings of this systematic review highlight several practical insights regarding the technical architecture of personal health monitoring (PHM) systems, particularly in terms of algorithm selection, data processing pipelines, and system deployment environments.

**1. Algorithm Selection Aligned with Data Type and System Goals**

The reviewed studies demonstrate that the choice of AI algorithm must be closely aligned with the nature of the input data and the intended system functionality. For example:

- Deep learning models, particularly CNNs and CNN-LSTM hybrids, as used in Studies 1, 6, 9, and 13, are best suited for tasks involving complex time-series physiological data such as ECG and PPG, where automated feature extraction is crucial.

- Traditional machine learning models such as SVM and Random Forest (Studies 3, 5, 10, 15) remain highly practical for behavioral classification, stress detection, or step activity recognition, especially in low-resource or mobile contexts, due to their lower computational cost and ease of deployment.

These insights support a modular design approach where algorithm selection is not one-size-fits-all, but tailored to the specific signal modality and computational constraints.

**2. Importance of Real-Time Processing and Edge AI**

Several studies (e.g., **Studies 2, 4, 7**) implemented their models on **edge devices** or in environments where **real-time feedback** was critical. This points to a growing need for:

- **On-device inference capabilities**, reducing reliance on cloud services.
- Lightweight model architectures optimized for memory and power efficiency (e.g., quantized neural networks or pruned ML models).

Such constraints must be integrated into the PHM system design process, particularly when targeting wearables, mobile health apps, or home-based monitoring devices.

To illustrate these considerations, the following table summarizes model types and their suitability based on system deployment level:

**Table 14: Algorithm Suitability Based on Deployment Context**

| Deployment Context | Preferred Algorithm Types | Key Considerations |
|---|---|---|
| Wearable / Edge Device | SVM, Random Forest, Shallow NN | Lightweight, interpretable, low latency |
| Mobile App | CNN, RNN, Hybrid ML-DL | Real-time processing, moderate complexity |
| Cloud-Based System | Deep CNN, Transformer-based DL | High computation power, rich contextual data handling |

**3. Integration of Multimodal Sensor Inputs**

Studies such as **Study 6 and 13** demonstrated that systems leveraging **multimodal data** (e.g., combining ECG, accelerometer, and temperature sensors) showed higher accuracy and robustness. Thus, practical system design should prioritize:

- Synchronization across sensors.
- Efficient **data fusion algorithms**, such as early fusion (at input level) or late fusion (at decision level).
- Handling of **missing or noisy data** through robust preprocessing and imputation methods.

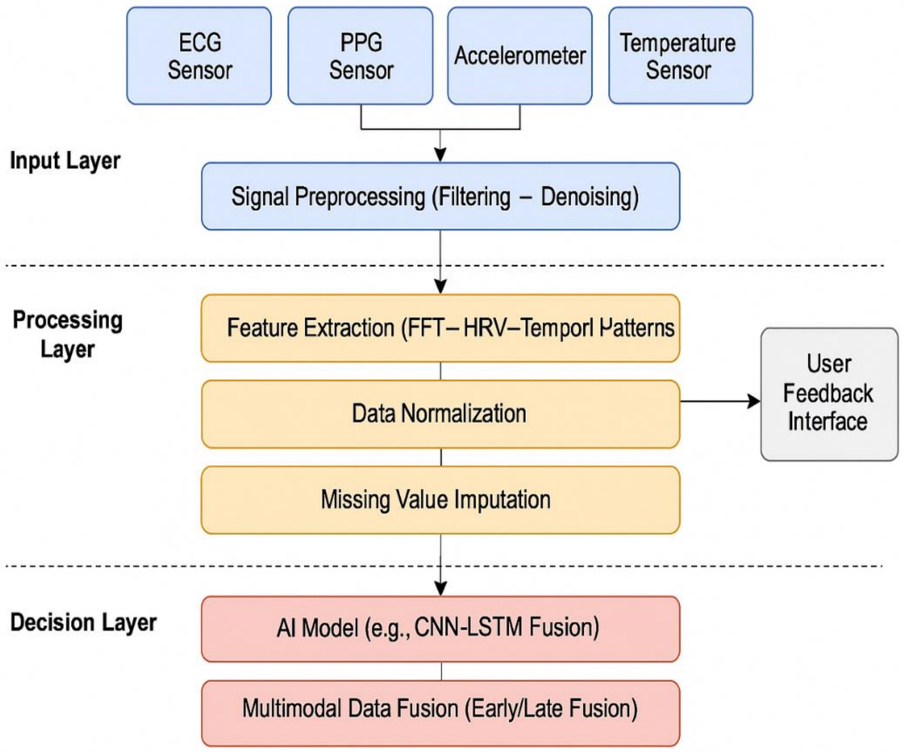


**Figure 11: Layered Architecture of a Personal Health Monitoring System**

**Source: Adapted from Nzomo & Moodley (2023), "Semantic Technologies in Sensor-Based Personal Health Monitoring Systems: A Systematic Mapping Study."**

**4. Model Evaluation for Practical Deployment**

Several reviewed studies lacked robust validation procedures beyond accuracy reporting. For system designers, practical deployment requires:

- Use of **cross-validation**, **confusion matrices**, and **statistical metrics** (e.g., F1-score, precision, recall).
- Inclusion of **confidence intervals** and **ROC curves** for risk-sensitive decisions.

Hence, PHM system developers must embed evaluation dashboards and logging mechanisms that monitor real-world model performance post-deployment—especially in adaptive or continuously learning systems.

**Second: User-Centric Design, Privacy, and System Adoption Considerations**

While algorithm performance and computational efficiency are essential, the real-world success of personal health monitoring (PHM) systems depends just as critically on user experience, data privacy, ethical design, and clinical integration. Insights from the reviewed studies, combined with best practices from the health informatics literature, suggest several key implications for PHM system developers and stakeholders.

**1. Privacy-by-Design and Data Governance**

A recurring concern in PHM systems is the sensitivity of health data, especially in continuous monitoring scenarios. As seen in Studies 2, 7, and 15, federated learning and on-device data processing offer promising privacy-preserving approaches. From a design perspective, this implies:

- Minimizing raw data transmission through local processing.
- Embedding consent management modules, enabling users to control data sharing.
- Ensuring compliance with frameworks like GDPR and HIPAA, particularly when integrating with external health databases or cloud services.

Designers must prioritize data minimization, transparent privacy policies, and auditability, especially for consumer-facing applications.

**2. User Experience and Interface Design**

The adoption and long-term use of PHM systems depend heavily on how users perceive and interact with the system. Despite the technical sophistication of many models, studies such as Study 5 and 10 noted issues of user disengagement due to complexity or lack of feedback. Key recommendations include:

- Providing real-time, interpretable feedback (e.g., via color-coded alerts or health scores).
- Designing adaptive interfaces that account for user preferences, health literacy, and accessibility.
- Reducing alert fatigue by implementing threshold-based or intelligent notification systems.

This process can be conceptualized as a user feedback loop involving data collection, local analysis, feedback display, optional cloud synchronization, and user action.

**3. Clinician Integration and System Trust**

Several studies (e.g., Study 6 and 13) emphasized the importance of systems that interface with healthcare professionals, allowing bidirectional data flow and collaborative decision-making. Practical design implications include:

- Interoperability with EHR systems (Electronic Health Records), through standardized APIs (e.g., HL7 FHIR).
- Generating clinician-friendly summaries, such as trend graphs or anomaly flags, to aid rapid assessment.
- Incorporating explainable AI modules that provide justification for predictions, aiding both user and clinician trust.

Lack of such mechanisms—observed in more technically focused studies—can severely limit real-world adoption, even for high-performing models.

**4. Equity, Personalization, and Inclusivity**

A final implication concerns design fairness and population inclusiveness. Several studies (notably Study 8 and 11) used datasets that lacked diversity in age, gender, or health status, a common critique in AI health research (Challen et al., 2019). PHM systems must therefore:

- Be trained and tested on representative datasets, ensuring equitable performance.
- Include personalization settings, such as baseline calibration or adaptive thresholds.
- Avoid bias amplification by maintaining transparency about dataset limitations and model assumptions.
- Failure to account for demographic and contextual diversity may reinforce health disparities rather than mitigate them.

# Chapter Five
# Conclusion and Future Work

## Conclusion and Future Work

The final chapter of this thesis is devoted to presenting the overall conclusion and outlining potential directions for future research. While the preceding chapters have provided a comprehensive review, analysis, and interpretation of artificial intelligence (AI) algorithms in Personal Health Monitoring (PHM) systems, it is essential to consolidate these findings into a coherent conclusion that reflects both the academic and practical significance of the study. Separating this chapter highlights the value of the outcomes achieved and emphasizes the study's contribution to knowledge in the field of AI-driven healthcare solutions.

This chapter is structured into two main sections. The first presents recommendations for future research, offering insights into methodological improvements, unexplored areas, and emerging opportunities for advancing the integration of AI into PHM. The second section delivers the conclusion, summarizing the thesis contributions, reaffirming the research aims, and reflecting on the broader implications of the findings for both scholars and practitioners. By doing so, this chapter not only closes the current study but also serves as a foundation for subsequent investigations and innovations in AI-powered health monitoring technologies.

## 5-1 Recommendations for Future Research

While the current systematic review provides a comprehensive synthesis of AI algorithms used in personal health monitoring (PHM), it also uncovers several methodological, technical, and practical gaps that future research must address to advance the field. The following recommendations are derived from critical analysis of the 15 selected studies and are organized to guide researchers across multiple domains, including computer science, health informatics, biomedical engineering, and clinical practice.

### 1. Expand Real-World Clinical Validation

A significant number of reviewed studies especially Studies 4, 8, and 11relied primarily on synthetic or publicly available datasets that lacked diversity, clinical realism, or temporal depth. While these datasets are valuable for preliminary validation, they do not adequately represent the complexities of real-world deployment.

To address this gap, future research should:

- Collaborate closely with healthcare institutions and clinics to obtain access to longitudinal, representative patient data collected in naturalistic settings.
- Design prospective clinical trials or pilot deployments that evaluate model performance in real hospital or home-care environments, involving varied populations and real-time conditions.
- Employ rigorous ground truth validation methods, including cross-referencing with electronic health records (EHRs) and conducting physician-confirmed outcome assessments.

Such steps are essential not only for improving model accuracy but also for ensuring external validity and real-world applicability, in line with the expectations of translational AI research (Esteva et al., 2021; Kelly et al., 2019).

**2. Prioritize Model Explainability and Transparency**

As AI systems become increasingly embedded in clinical decision-making processes, their acceptance by healthcare professionals hinges on the clarity and interpretability of their outputs. Despite their high accuracy, many models especially deep learning architectures like CNNs and LSTMs function as black boxes, which poses challenges for clinical trust and regulatory approval.

Among the reviewed studies, only a limited number (e.g., Study 6 and 15) included mechanisms for explainability, such as visualization layers or interpretable rule sets. To bridge this critical gap, future research should:

- Incorporate state-of-the-art explainable AI (XAI) methods, such as SHAP (SHapley Additive exPlanations), LIME (Local Interpretable Model-agnostic Explanations), Grad-CAM, or attention-based saliency maps, to clarify how predictions are made.
- Ensure transparency by publishing open-source code, training datasets (where ethically permissible), and full documentation of model architecture and decision logic.
- Explicitly report model limitations, biases, edge cases, and decision thresholds in a standardized, peer-auditable format. Enhancing explainability not only fosters clinician confidence but also aligns with growing international expectations for transparent and responsible AI in healthcare (Amann et al., 2020; European Commission, 2019).

### 3. Develop Lightweight and Personalized Models

As personal health monitoring (PHM) increasingly shifts toward wearable and mobile platforms, the need for computationally efficient and adaptive AI models becomes paramount. Traditional deep learning architectures, while powerful, are often too resource-intensive for deployment on low-power edge devices. Moreover, static models fail to account for the diversity in individual health patterns. To address these challenges, future research should:

- Explore model optimization techniques such as pruning, quantization, and knowledge distillation to reduce computational demands without compromising performance.
- Implement personalization strategies through meta-learning or transfer learning, allowing models to adapt to user-specific physiological baselines and behavioral trends over time.
- Leverage on-device federated learning to continuously update models while maintaining user privacy, especially in contexts where cloud-based solutions are impractical or ethically sensitive.

These directions will support the development of inclusive and scalable PHM systems that offer tailored insights without requiring constant connectivity or high-end hardware.

### 4. Standardize Benchmarking and Reporting

A recurring limitation identified in the reviewed studies is the absence of consistent benchmarking protocols and reporting standards. Performance metrics, dataset descriptions, and evaluation methods varied widely, impeding the reproducibility and comparability of findings. These inconsistencies reduce the ability to draw reliable conclusions across studies and slow collective progress in the field. To address this issue, future research should:

- Adopt common benchmarking datasets and evaluation frameworks where feasible, particularly for well-studied PHM tasks such as ECG classification or activity recognition.
- Provide detailed documentation of model architecture, training configurations (e.g., learning rates, epochs), and computational environments used during experimentation.
- Report not only accuracy but also additional statistical metrics such as confidence intervals, p-values, and clinical utility measures that reflect practical relevance.

## 5-2 Conclusion

This systematic review set out to explore and evaluate the application of artificial intelligence algorithms in personal health monitoring (PHM) systems, focusing on algorithm types, performance metrics, ethical considerations, and practical deployment aspects. By analyzing fifteen recent and diverse studies, the review offers a detailed and evidence-based understanding of the current landscape of AI-driven PHM.

The findings reveal that while AI models particularly deep learning and hybrid architectures have shown strong promise in tasks such as anomaly detection, stress monitoring, and physical activity recognition, several challenges persist. These include limited real-world validation, inconsistent performance reporting, underdeveloped user experience considerations, and variable attention to data privacy and ethical governance.

From a technical standpoint, the review underscores the importance of selecting algorithms that align with system goals, resource constraints, and input signal characteristics. It also emphasizes the need for efficient processing models for wearable and mobile deployments, as well as the potential of multimodal data fusion to enhance system robustness.

Beyond the technical domain, the review highlights the necessity of human-centered design, clinician engagement, and equitable model training as critical factors for ensuring that PHM systems are not only functional but also trusted and adopted. The practical implications drawn from the analysis provide developers, researchers, and healthcare stakeholders with a comprehensive guide for designing next-generation PHM systems.

Finally, the recommendations outlined for future research—centered on clinical validation, explainability, personalization, ethical rigor, and standardization serve as a roadmap for addressing the gaps identified in the current literature and for guiding future innovation in this rapidly evolving field.

In conclusion, AI-powered PHM systems hold immense potential to transform healthcare delivery, especially in the domains of preventive care, chronic disease management, and real-time patient monitoring. However, unlocking this potential requires not only advanced technical innovation but also careful integration with ethical principles, clinical relevance, and user-centered design.

# References


1. Alam, T., Mehmood, R., Katib, I., & Albeshri, A. (2018). Data fusion and IoT for smart ubiquitous environments: A survey. *IEEE Access*, 6, 21075–21093. https://doi.org/10.1109/ACCESS.2018.2825282
2. Altun, K., Barshan, B., & Tuncel, O. (2010). Comparative study on classifying human activities with miniature inertial and magnetic sensors. *Pattern Recognition*, 43(10), 3605–3620. https://doi.org/10.1016/j.patcog.2010.04.019
3. Amann, J., Blasimme, A., Vayena, E., Frey, D., & Madai, V. I. (2020). Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Medical Informatics and Decision Making*, 20(1), 310. https://doi.org/10.1186/s12911-020-01332-6
4. Amann, J., Vayena, E., & Frey, D. (2022). Design thinking for artificial intelligence in health care: The case of clinical decision support systems. *Journal of Medical Internet Research*, 24(3), e25325. https://doi.org/10.2196/25325
5. Banaee, H., Leu, M. C., & Ahmed, M. U. (2013). Data mining for wearable sensors in health monitoring systems: A review of recent trends and challenges. *Sensors*, 13(12), 17472–17500. https://doi.org/10.3390/s131217472
6. Chen, M., Ma, Y., Li, Y., Wu, D., Zhang, Y., & Youn, C. H. (2017). Wearable 2.0: Enabling human-cloud integration in next generation healthcare systems. *IEEE Communications Magazine*, 55(1), 54–61. https://doi.org/10.1109/MCOM.2017.1600331CM
7. Chung, A. E., et al. (2019). Patient engagement strategies in PHM: How real-time feedback changes behavior. *Journal of Medical Internet Research*, 21(7), e12533. https://doi.org/10.2196/12533
8. Collins, G. S., Reitsma, J. B., Altman, D. G., & Moons, K. G. M. (2015). Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. *Annals of Internal Medicine*, 162(1), 55–63. https://doi.org/10.7326/M14-0697
9. Dai, L., Wu, L., Li, H., Cai, C., Wu, Q., Kong, H., ... & Jia, W. (2021). A deep learning system for detecting diabetic retinopathy across the disease spectrum. Nature communications, 12(1), 3242.
10. Dey, N., Ashour, A. S., & Balas, V. E. (2021). *Smart Medical Data Sensing and IoT Systems Design in Healthcare*. Springer.

11. Esteva, A., Kuprel, B., Novoa, R. A., et al. (2019). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115–118. https://doi.org/10.1038/nature21056
12. Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., ... & Dean, J. (2021). A guide to deep learning in healthcare. *Nature Medicine*, 27(4), 758–765. https://doi.org/10.1038/s41591-021-01301-6
13. European Commission. (2019). Ethics Guidelines for Trustworthy AI. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai
14. FDA. (2021). *Artificial Intelligence and Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD)*. U.S. Food and Drug Administration. https://www.fda.gov/medical-devices
15. Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., ... & Webster, D. R. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. jama, 316(22), 2402-2410.
16. Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and Machines*, 30(1), 99–120. https://doi.org/10.1007/s11023-020-09517-8
17. Hannun, A. Y., Rajpurkar, P., Haghpanahi, M., et al. (2019). Cardiologist-level arrhythmia detection with deep learning. *Nature Medicine*, 25(1), 65–69. https://doi.org/10.1038/s41591-018-0268-3
18. Islam, S. M. R., Kwak, D., Kabir, M. H., Hossain, M., & Kwak, K. S. (2015). The Internet of Things for health care: A comprehensive survey. *IEEE Access*, 3, 678–708. https://doi.org/10.1109/ACCESS.2015.2437951
19. Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine*, 17(1), 195. https://doi.org/10.1186/s12916-019-1426-2
20. Korhonen, I., Välimäki, E., & Pärkkä, J. (2003). Wireless sensors for remote healthcare monitoring in home and hospital environments. *Computers in Cardiology*, 30, 373–376.
21. Li, X., Zhang, D., & Wang, Y. (2021). AI applications in PHM: A review of classification algorithms and use cases. *Healthcare Technology Letters*, 8(1), 23–30. https://doi.org/10.1049/htl2.12009
22. Lipton, Z. C., Kale, D. C., Elkan, C., & Wetzel, R. (2016). Learning to diagnose with LSTM recurrent neural networks. *arXiv preprint* arXiv:1511.03677

23. Litjens, G., Kooi, T., Bejnordi, B. E., et al. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60–88. https://doi.org/10.1016/j.media.2017.07.005
24. Liu, N. T., Holcomb, J. B., & Wade, C. E. (2020). Data science in health monitoring: A review of AI methods and applications. *IEEE Reviews in Biomedical Engineering*, 13, 20–34. https://doi.org/10.1109/RBME.2019.2917327
25. Luxton, D. D., McCann, R. A., Bush, N. E., Mishkind, M. C., & Reger, G. M. (2011). mHealth for mental health: Integrating smartphone technology in behavioral healthcare. *Professional Psychology: Research and Practice*, 42(6), 505–512. https://doi.org/10.1037/a0024485
26. Marsch, L. A., & Gustafson, D. H. (2013). *Behavioral Healthcare and Technology: Using Science-Based Innovations to Transform Practice*. Oxford University Press.
27. Mishra, T., Wang, M., Metwally, A. A., et al. (2020). Pre-symptomatic detection of COVID-19 from smartwatch data. *Nature Biomedical Engineering*, 4(12), 1208–1220. https://doi.org/10.1038/s41551-020-00640-6
28. Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 1–21. https://doi.org/10.1177/2053951716679679
29. Muoio, D. (2020). FDA clears wearable ECG patch for long-term monitoring. *MobiHealthNews*. https://www.mobihealthnews.com
30. Nguyen, K. T., Medjaher, K., & Tran, D. T. (2023). A review of artificial intelligence methods for engineering prognostics and health management with implementation guidelines. Artificial Intelligence Review, 56(4), 3659-3709.
31. Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447–453. https://doi.org/10.1126/science.aax2342
32. Osowski, S., Hoai, L. T., & Markiewicz, T. (2004). Support vector machine-based expert system for reliable heartbeat recognition. *IEEE Transactions on Biomedical Engineering*, 51(4), 582–589. https://doi.org/10.1109/TBME.2004.824132
33. Palakurti, N. R. (2023). AI-Driven Personal Health Monitoring Devices: Trends and Future Directions. ESP Journal of Engineering & Technology Advancements, 3(3), 41-51.

34. Pantelopoulos, A., & Bourbakis, N. G. (2010). A survey on wearable sensor-based systems for health monitoring and prognosis. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 40(1), 1–12. https://doi.org/10.1109/TSMCC.2009.2032660
35. Patel, S., Park, H., Bonato, P., Chan, L., & Rodgers, M. (2012). A review of wearable sensors and systems with application in rehabilitation. *Journal of NeuroEngineering and Rehabilitation*, 9(1), 21. https://doi.org/10.1186/1743-0003-9-21
36. Piwek, L., Ellis, D. A., Andrews, S., & Joinson, A. (2016). The rise of consumer health wearables: Promises and barriers. *PLOS Medicine*, 13(2), e1001953. https://doi.org/10.1371/journal.pmed.1001953
37. Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine*, 380(14), 1347–1358. https://doi.org/10.1056/NEJMra1814259
38. Rajkomar, A., et al. (2018). Scalable and accurate deep learning with electronic health records. *NPJ Digital Medicine*, 1, 18. https://doi.org/10.1038/s41746-018-0029-1
39. Rajpurkar, P., Hannun, A. Y., Haghpanahi, M., Bourn, C., & Ng, A. Y. (2017). Cardiologist-level arrhythmia detection with convolutional neural networks. *arXiv preprint* arXiv:1707.01836
40. Ravi, D., Wong, C., Lo, B., & Yang, G. Z. (2017). Deep learning for health informatics. *IEEE Journal of Biomedical and Health Informatics*, 21(1), 4–21. https://doi.org/10.1109/JBHI.2016.2636665
41. Rehman, A. U., et al. (2020). Wearable sensors and AI for digital health: Challenges and opportunities. *Sensors*, 20(18), 5379. https://doi.org/10.3390/s20185379
42. Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H. R., Albarqouni, S., ... & Cardoso, M. J. (2020). The future of digital health with federated learning. *NPJ Digital Medicine*, 3(1), 119. https://doi.org/10.1038/s41746-020-00323-1
43. Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x
44. Sheller, M. J., Edwards, B., Reina, G. A., Martin, J., Pati, S., & Bakas, S. (2020). Federated learning in medicine: facilitating multi-institutional collaborations

without sharing patient data. *Scientific Reports*, 10(1), 12598. https://doi.org/10.1038/s41598-020-69250-1

45. Shen, C., Tan, M., Wu, X., & Guan, Y. (2021). Smart health: Concepts, trends, and challenges. *IEEE Internet of Things Journal*, 8(8), 6197–6210. https://doi.org/10.1109/JIOT.2020.3025784
46. Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: Vision and challenges. *IEEE Internet of Things Journal*, 3(5), 637–646. https://doi.org/10.1109/JIOT.2016.2579198
47. Shickel, B., Tighe, P. J., Bihorac, A., & Rashidi, P. (2018). Deep EHR: A survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. *IEEE Journal of Biomedical and Health Informatics*, 22(5), 1589–1604. https://doi.org/10.1109/JBHI.2017.2767063
48. Swan, M. (2012). Sensor mania! The internet of things, wearable computing, objective metrics, and the quantified self 2.0. *Journal of Sensor and Actuator Networks*, 1(3), 217–253. https://doi.org/10.3390/jsan1030217
49. Sze, V., Chen, Y. H., Yang, T. J., & Emer, J. S. (2017). Efficient processing of deep neural networks: A tutorial and survey. *Proceedings of the IEEE*, 105(12), 2295–2329. https://doi.org/10.1109/JPROC.2017.2761740
50. Ting, D. S. W., Pasquale, L. R., Peng, L., et al. (2020). Artificial intelligence and deep learning in ophthalmology. *British Journal of Ophthalmology*, 104(2), 167–175. https://doi.org/10.1136/bjophthalmol-2019-314577
51. Tjoa, E., & Guan, C. (2020). A survey on explainable artificial intelligence (XAI): Toward medical XAI. *IEEE Transactions on Neural Networks and Learning Systems*, 32(11), 4793–4813. https://doi.org/10.1109/TNNLS.2020.3027314
52. Topol, E. (2015). *The Patient Will See You Now: The Future of Medicine is in Your Hands*. Basic Books.
53. Topol, E. (2019). *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. Basic Books.
54. Topol, E. (2019). *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. Basic Books.
55. Tschider, C. A. (2021). International data transfers under the GDPR: Regulatory responses and global data governance. *Chicago Journal of International Law*, 22(1), 69–110.

56. Wang, Y., Qian, Y., & Zhang, Y. (2019). Multimodal emotion recognition using deep learning and decision-level fusion. *Information Fusion*, 49, 46–54. https://doi.org/10.1016/j.inffus.2018.09.001
57. Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology*, 10(2), 1–19. https://doi.org/10.1145/3298981
58. Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology*, 10(2), 1–19. https://doi.org/10.1145/3298981
59. Zhou, T., et al. (2021). Gradient boosting decision tree-based wearable sensor data classification for health monitoring. *Sensors*, 21(4), 1217. https://doi.org/10.3390/s21041217
60. Zhou, X., et al. (2019). Anomaly detection in ECG time series: A deep learning approach. *IEEE Access*, 7, 67775–67785. https://doi.org/10.1109/ACCESS.2019.2917760
61. Zhou, X., Wang, J., & Hu, B. (2019). Decision tree-based anomaly detection algorithms in wearable health monitoring systems. *Sensors*, 19(6), 1495. https://doi.org/10.3390/s19061495
62. Nie, W., Lu, X., Wang, Y., Zhang, Y., Zhang, Y., & Al-Qaisi, F. (2025). Vascular age derived from photoplethysmography signals using deep learning: A new biomarker for cardiovascular risk. *arXiv preprint*, arXiv:2502.12990. https://arxiv.org/abs/2502.12990
63. Shaik, M. N., Alloghani, M., Al-Jumeily, D., Aljaaf, A. J., Baker, T., Hussain, A., & Mohamed, A. (2023). A comprehensive review on artificial intelligence in remote patient monitoring systems. *arXiv preprint*, arXiv:2301.10009. https://arxiv.org/abs/2301.10009
64. Ye, J., Liu, Z., Pan, J., Zhang, S., Zheng, Y., & Li, G. (2024). DActAHM: Dynamic activity-aware health monitoring via reinforcement learning and SlowFast models. *arXiv preprint*, arXiv:2401.10794. https://arxiv.org/abs/2401.10794
65. Nzomo, M., & Moodley, D. (2023). Semantic Technologies in Sensor-Based Personal Health Monitoring Systems: A Systematic Mapping Study. Preprint.