STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **15**, February 2026, pp 1398–1415.
Published online in International Academic Press (www.IAPress.org)

Contextual Wellness Recommender System: Leveraging IoT Data for contextual Healthcare Recommendations

AIT ELBAZ Aicha [1], STITINI Oumaima [2,3], KALOUN Soulaimane [2]

[1] *Mohammed V University, Mohammadia School of Engineers, LERMA-IFE Laboratory, Rabat, Morocco*
[2] *Cadi Ayyad University, Faculty of Sciences and Technologies, L2IS Laboratory, Marrakesh, Morocco*
[3] *Cadi Ayyad University, Higher Normal School, Department of Computer Science, Marrakesh, Morocco*

Abstract Since their introduction, personalized recommendation systems have experienced remarkable evolution, aiming to provide recommendations corresponding to the needs and preferences of users, particularly in the field of healthcare. However, these systems have encountered limitations, particularly regarding the dynamic nature of users health needs. Contextual recommender systems take this dynamic into account and use users contextual data to generate personalized recommendations tailored to their needs. In this context ,Internet of Things (IoT) technologies assume a pivotal role, enabling remote monitoring of various health aspects, facilitating proactive healthcare interventions, and crafting personalized treatment plans. The data collected by IoT devices serves as a valuable resource, enhancing the effectiveness of personalized recommendations.To address this challenge, we propose a novel approach named “Contextual Wellness Recommender System”, a methodology that fully exploits the data collected by several connected health devices. Our methodology relies on the use of advanced machine learning and data analysis techniques to intelligently integrate contextual information into the recommendation process.Using machine learning algorithms, we will train two models to recognize patterns and correlations in IoT sensors data and other contextual factors. The used data includes users physical activity,demographics, lifestyle habits and vital signs. By taking into account these multiple dimensions of data, our models will be able to generate personalized recommendations that will allow users to proactively take care of their health.With an accuracy of over 90% on Model_A and more than 80% on Model_B on both training and validation data,our proposed approach stands out for its use of several and diverse connected health devices to generate recommendations. This innovative approach increases efficiency, personalization and adaptability by adapting to different individuals health conditions and fully exploiting their contextual data.

Keywords Context-aware recommendation systems, IoT, Contextual information, Recommendation Systems

DOI: 10.19139/soic-2310-5070-2766

1. Introduction

Recommender systems (RSs) have become integral tools across various domains, enhancing user experiences by providing suggestions to users based on their preferences and behaviors. Over time, these systems have evolved from traditional models to sophisticated personalized approaches that analyze detailed user data [1]. Despite their advancements, traditional recommender systems face significant limitations, particularly in adapting to the dynamic nature of user profiles. As user preferences can change due to various factors, recommendations based solely on historical data may become less relevant over time. To address this challenge, context-aware recommender systems (CARS) have emerged, offering recommendations that consider the contextual information of users, such as their current location, time of day, social environment, and specific activities[2]. By integrating these contextual factors, CARS aim to provide more accurate and timely recommendations that align with the user’s immediate needs and circumstances.

*Correspondence to: AIT ELBAZ Aicha (Email: aichaaitelbaz2000@gmail.com). Mohammed V University, Mohammadia School of Engineers, LERMA-IFE Laboratory, Rabat, Morocco.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

However, the effectiveness of CARS is often hindered by issues related to contextual data accessibility and quality[3]. Obtaining real-time, relevant contextual data can be challenging, and ensuring the reliability and accuracy of this data is crucial for delivering effective recommendations. In this paper, we explore the evolution of recommender systems towards personalized approaches and examine the inherent limitations of traditional methods. We then introduce context-aware recommender systems as a promising solution to address these limitations, highlighting the challenges associated with contextual data.

The Internet of Things (IoT) emerges as a promising source of contextual data, providing real-time and relevant information that can significantly enhance the effectiveness of recommender systems[5]. Our research focuses on applying this approach in the field of healthcare, where personalized recommendations can significantly impact patient outcomes. By integrating data from connected health devices, we aim to build a contextual recommender system that provides accurate and personalized healthcare recommendations. This integration of Internet of Things (IoT) technology with context-aware recommendation methods represents a novel approach to enhancing healthcare services, ultimately improving patient care and satisfaction.

This paper is structured as follows: Section 2 introduces Personalized Recommendation Systems, while Section 3 provides an overview of Traditional Recommender Systems and their limitations. The Problematic addressed is outlined in Section 4. Section 5 presents a Background of context-aware systems and explores the role of combining the Internet of Things (IoT) with contextual approaches. Section 6 outlines our Proposed Approach, followed by Experiments and Results in Section 7, and a final Discussion in Section 8.

2. Personalized Recommendation Systems

Recommender systems have undergone significant evolution since their inception, transforming from rudimentary tools to sophisticated, personalized systems. The earliest recommender systems were non-personalized, relying on generic lists of bestsellers and popular items to guide users [6]. These systems operated on the assumption that the popularity of an item would translate to user interest, which, while effective to an extent, lacked the nuance needed to cater to individual preferences. In these early stages, the focus was on leveraging aggregate data. For instance, a bookstore might display a list of top-selling books, assuming that new customers would find these suggestions appealing based on their widespread popularity. Similarly, online platforms would provide sections like "Most Viewed" or "Top Rated" items. While this approach provided a basic level of guidance, it was inherently limited. It didn't account for the diverse tastes and preferences of individual users, often leading to a less satisfying user experience. The advent of personalized recommender systems marked a significant leap forward [7]. These systems began to incorporate data about individual users, such as their past behavior, preferences, and interactions. By analyzing this data, personalized recommender systems can make tailored suggestions that resonate more closely with the user's unique tastes. This shift was driven by advancements in data collection and processing, as well as the development of sophisticated algorithms capable of making sense of vast amounts of user data.

3. Typology of Traditional Recommender Systems

In the realm of traditional recommender systems, the most widely used approaches include collaborative filtering, content-based filtering, and hybrid methods [8]. These methods represented in figure 1, form the cornerstone of recommendation technologies and have been extensively developed and applied across various domains.

3.1. Collaborative Filtering

Collaborative Filtering operates on the principle that users who have previously shown similar preferences are likely to continue agreeing on future items as well [1]. It uses user-item interactions to find patterns and make recommendations.Netflix's recommendation engine primarily uses collaborative filtering to suggest movies and TV shows [9]. If two users have a high overlap in their viewing histories, Netflix will recommend shows that one user has watched but the other has not yet seen.

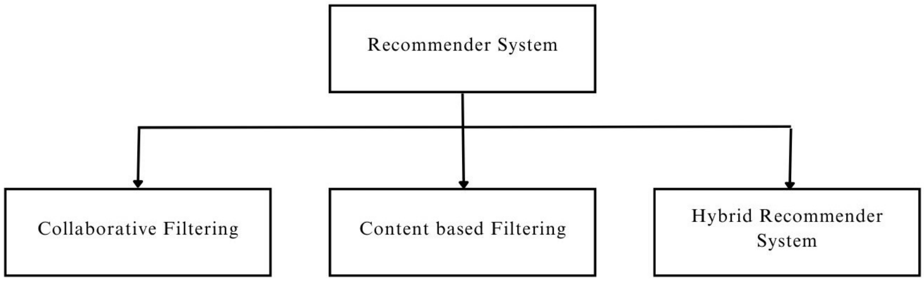

Figure 1. Traditional Recommender Systems Typology

We can distinguish two major categories of collaborative filtering: User-based and Item-based. User-based collaborative filtering recommends items based on the similarity between users, while Item-based collaborative filtering recommends items by examining the similarities among various items [4].

3.2. *Content-Based Filtering*

Content-Based Filtering recommends items by analyzing the content of items and matching them with user preferences [10]. This approach relies on the characteristics or features of items and the user's previous interactions with these features. It develops a user profile by considering the characteristics of items that the user has interacted with in the past [11]. For example, consider a book recommendation system in a local library. If a user frequently borrows books in the mystery and thriller genres, the system will analyze the content of these books such as themes, keywords, and author styles and create a profile for the user. Based on this profile, the system will recommend other mystery and thriller books that have similar characteristics, even if the user has never borrowed or shown interest in these specific titles before.

3.3. *Hybrid Recommender Systems*

Classical recommendation methods often face several limitations, such as the cold start problem, where new users or items lack sufficient data for effective recommendations [12]. They can also suffer from overspecialization, where recommendations become too narrow and fail to expose users to diverse options. Moreover, purely collaborative or content-based approaches may not fully capture user preferences or item characteristics, leading to less accurate suggestions.

Hybrid recommender systems tackle these issues by integrating various recommendation methods, allowing them to take advantage of each technique's strengths while minimizing their individual shortcomings. For instance, by integrating collaborative filtering, which relies on user interactions, with content-based filtering, which analyzes item attributes, hybrid systems can provide more comprehensive insights [13]. This integration allows them to deliver personalized recommendations that are not only relevant but also diverse, enhancing user experience.

Additionally, hybrid systems can utilize advanced methods like matrix factorization and deep learning to further refine their recommendations [14].

4. Problematic

Despite significant advancements in creating personalized user experiences, traditional recommender systems still face inherent limitations. A key issue lies in their inability to adapt to the dynamic nature of user profiles and preferences. User preferences are not static ; they evolve over time and can be influenced by a variety of factors[18], such as changing tastes, contextual situations, and external influences. However, traditional systems rely heavily on historical user data, which often results in outdated and less relevant recommendations.

One of the main limitations of these systems, as described by Adomavicius and Tuzhilin[1], is their limited approach to user profiling. They tend to operate under the assumption that a user's preferences remain constant over time, an assumption that rarely holds true. For example, a user might prefer action movies when alone but opt for family-friendly content when watching with others, or a user may initially show interest in a particular genre of movies or a specific type of product, but over time, these preferences may shift due to new experiences, trends, or even situational factors [12].

Traditional systems continue to generate recommendations based on outdated data, overlooking these evolving preferences. Moreover, traditional recommender systems often fail to incorporate additional factors that influence how users interact with the system. By ignoring these factors, such systems miss opportunities to provide more timely and contextually relevant recommendations, leading to a less optimal user experience.

4.1. Research Questions

1. How to effectively manage the fluctuating nature of user profiles to provide more personalized and relevant recommendations?

The disconnect between traditional recommender systems and the evolving nature of user behavior highlights the need for new strategies that ensure recommendations remain timely and relevant. One key challenge is managing the fluctuating nature of user profiles, and addressing this requires systems that can adapt to these dynamic user profiles and provide more personalized recommendations.

In section 5, we introduce Context-Aware Recommender Systems (CARS) as a solution to better handle dynamic user profiles. CARS leverages contextual information to provide recommendations that are not only personalized but also responsive to changes in user behavior and external factors.

In addition to the challenge of user profiling, another critical research question we're going to cover during this study is:

2. How can IoT integration improve the drawbacks of contextual recommendation systems?

The integration of external data sources such as real-time data from IoT and connected devices offers new opportunities to improve contextual data accessibility. By incorporating these technologies, we can transform the relevance and accuracy of recommendations across various fields.

This research questions leads us to a deeper exploration of the importance of both CARS and IoT technologies, as they present powerful tools for enhancing the contextual relevance and adaptability of recommendations, making them more responsive to real-time factors and user environments.

4.2. Business Problematic: Healthcare Domain

1. How can personalized recommendation systems integrate connected health data to improve quality of care and patient engagement?

Recommender systems (RSs) initially rose to prominence in the e-commerce sector, where they revolutionized the way companies deliver personalized experiences to users by offering tailored product suggestions based on their preferences and behaviors. As their success in e-commerce grew, recommender systems began to emerge in various other fields, including media, entertainment, and even education. However, among these sectors, healthcare stands out as one of the most critical and sensitive areas for recommender systems [15].

In healthcare, RSs are not merely tools for enhancing user experiences, they have the potential to directly impact patient care, and in some cases even save lives. The performance of recommender systems in this domain is, therefore, not just beneficial, it is essential. These systems must deliver highly personalized recommendations that account for the unique and sensitive nature of individual health needs.

With the rise of connected health technologies, such as wearable devices, mobile health apps, and remote monitoring systems, there is now an unprecedented opportunity to incorporate real-time, patient-specific data into the recommendation process [16] [17]. This integration of connected health data can further enhance the accuracy, relevance, and responsiveness of healthcare recommendations, ensuring they align with the ever-changing needs of patients.

5. Background

5.1. Context-Aware Recommender Systems

Context-aware recommender systems (CARS) is one of the important innovations in the field of recommendation technologies. These systems aim to enhance the accuracy of recommendations by integrating contextual information into the recommendation system. This means that the recommender system as the figure 2 shows, will take as input not only information about users and items but also a third dimension of data represented by the context [2]. Unlike traditional recommender systems that rely solely on historical user data, CARS take into account various contextual factors that can influence user preferences and decision-making at any given moment [3].

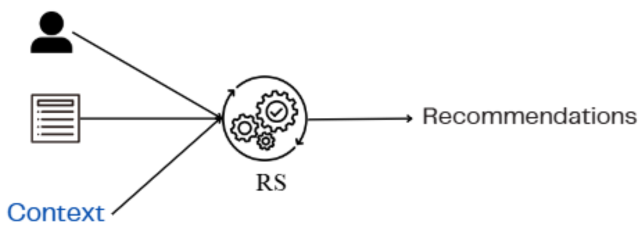

Figure 2. Context Aware Recommender System

5.1.1. How can context be defined? Context, within the scope of context-aware recommender systems, refers to any information that defines the situation of an entity relevant to the interaction between the user and the system. This information reflects better the user's current state and circumstances, which play a crucial role in shaping his preferences and decisions.

Context can be broadly categorized into several dimensions [19], take a look on figure 3. each of this dimentions plays a crucial role in understanding user circumstances and enhancing the accuracy and relevance of recommendations.

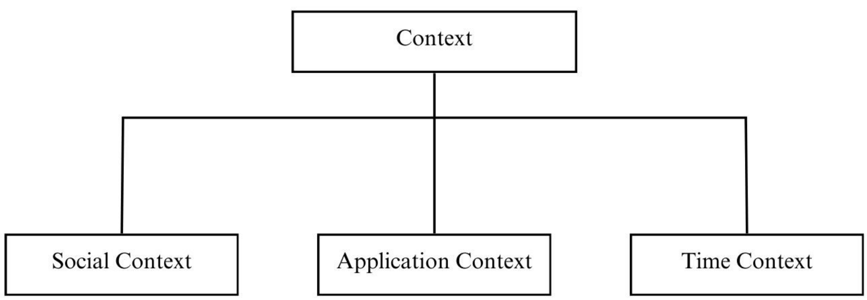

Figure 3. Context Dimensions.

- Social Context :
  Social context refers to the influence of social interactions and relationships on user preferences and behavior. This includes information about the user's social network, group memberships, and social activities.For example, a music streaming service might use social context to enhance its recommendations: If a user

frequently interacts with friends who listen to jazz music, the service may recommend jazz playlists or artists to that user, even if the user wasn't interested about jazz in the past.

- Application Context:
  Application context involves the specific circumstances under which a user interacts with an application. This can include the type of device being used, the user's interaction history with the application, and the specific features or services being utilized.
- Time Context:
  Time context encompasses temporal factors that can influence user preferences and decisions. This includes the time of day, day of the week, season, or even specific events and holidays.

5.2. *Integrating Context into Recommender Systems*

Incorporating context into recommender systems marks a significant step forward in the field, enhancing the relevance and personalization of recommendations by taking into account different contextual factors [3]. The classification of context incorporation in recommender systems depends on how contextual data is utilized and integrated within the recommendation process. Here, we explore different methods for incorporating context, focusing on context-driven search and request approaches, and contextual preference elicitation and estimation techniques [20].

5.2.1. Context-Driven Search and Request Approach This approach utilizes contextual information gathered from user inputs or environmental sensors to retrieve pertinent data [2]. Once collected, this information is used to query a relevant database to find matches that align with the user's query. For instance, a location-based service might utilize GPS data to suggest nearby restaurants. By leveraging real-time contextual information, this method aims to offer recommendations that are directly relevant to the user's current circumstances

5.2.2. Contextual Preference Elicitation and Estimation This method seeks to understand user preferences across various contexts using established recommendation techniques or machine learning algorithms. It encompasses three fundamental approaches:

- Pre-filtering: Pre-filtering techniques involve integrating contextual information with input data before generating a recommendation list. This method aims to simplify the multidimensional matrix into a user item matrix suitable for traditional recommendation algorithms. However, its accuracy may be compromised when contextual extraction is challenging or when insufficient data exists on the user's preferences, resulting in less precise recommendations.
- Post-filtering: Post-filtering methods initially disregard contextual details when creating a recommendation list. Subsequently, recommendations are tailored to each user based on specific contextual factors. This adjustment involves either sorting recommendations according to contextual relevance or filtering out irrelevant suggestions. For instance, an e-commerce platform might initially recommend products based on purchase history and then reorder them based on current browsing context, such as holiday shopping preferences.
- Contextual Modeling: Contextual modeling integrates contextual information directly into the recommendation process. These approaches utilize predictive models and heuristic strategies to construct recommendation systems that account for multiple dimensions of user behavior and preferences. By embedding context into the recommendation algorithm, these methods often yield more sophisticated and accurate suggestions

5.3. *Limitations of Context-Aware Recommender Systems*

Context-aware recommender systems (CARS) have significantly enhanced personalization by integrating contextual information into the recommendation process. Nonetheless, these systems encounter various challenges, particularly concerning the accessibility and quality of contextual data compared to traditional approaches [21].

- **Accessibility of Contextual Data:** One of the primary challenges of context-aware recommender systems is the accessibility of real-time and relevant contextual data. Traditional methods of acquiring contextual information, such as relying on user inputs, may be limited in their ability to provide continuous and comprehensive data streams. For example, manually entered user preferences may not capture the dynamic and evolving context of user interactions effectively. This limitation can hinder the system's ability to deliver timely and accurate recommendations that cater to the user's current needs and preferences.
- **Quality of Contextual Data:** Even when contextual data is accessible, its quality remains a major challenge for the effectiveness of context-aware recommender systems. Traditional approaches to integrating contextual data often struggle with ensuring data quality across various dimensions. Issues such as data noise, incomplete information, or outdated data can undermine the system's ability to generate relevant recommendations.

Surmounting these limitations involves exploring alternative methods for acquiring and enhancing contextual data quality. Integrating diverse data sources to dynamically adjust to evolving user contexts can improve the relevance and accuracy of recommendations. By adopting such strategies, context-aware recommender systems can mitigate this challenges, thereby enhancing their capability to deliver personalized and timely recommendations aligned with users current needs and preferences.

5.4. Internet of Things

The Internet of Things (IoT) encompasses a network of devices embedded with sensors, software, and other technologies, allowing them to gather and exchange data via the internet [22]. This network includes a wide array of devices, from everyday items like wearables to industrial machinery and vehicles.
The core concept of IoT is to facilitate seamless communication and data exchange between physical objects and digital systems, enabling automation, monitoring, and control across numerous fields.As illustrated in figure 4, IoT applications span diverse domains such as healthcare, smart homes, industrial IoT, smart cities, agriculture, and transportation. Each of these areas utilizes IoT technology to enhance efficiency and effectiveness in various processes.

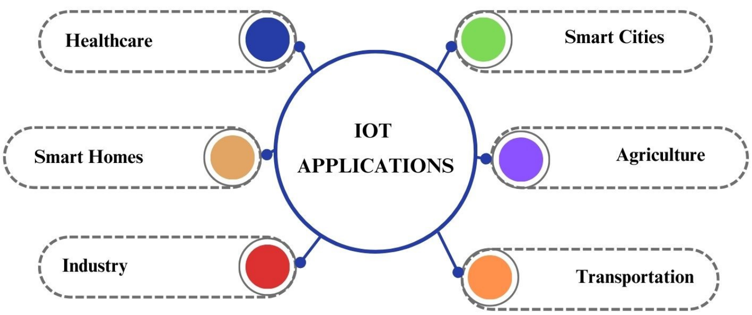

Figure 4. Applications of Internet of Things (IoT).

5.5. Combining IoT with Contextual Approach

Integrating the Internet of Things (IoT) with context-aware recommender systems offers a promising solution to address the limitations related to the accessibility and quality of contextual data . IoT devices, with their capability to gather and transmit real-time data from various sources, can provide a rich, continuous stream of contextual information that enhances the effectiveness of recommendations [23]. IoT devices, such as smartwatches, fitness trackers, smart home appliances, and environmental sensors, can continuously capture data about the user's current state and surroundings.By leveraging these IoT devices, context-aware recommender systems can access a wealth of real-time contextual data that would be difficult to obtain through traditional means. This continuous data flow ensures that recommendations are based on the most current and relevant information, improving their timeliness and accuracy .

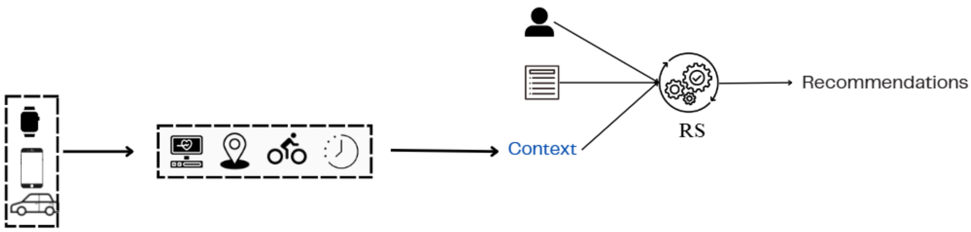

Figure 5. Integration of IoT Devices with Context-Aware Recommender Systems

6. Proposed approach

Recommender systems play a pivotal role across various domains, with healthcare emerging as one of the most critical areas of application. In healthcare, recommender systems serve as valuable tools on websites and platforms that recommend healthcare services or provide essential information to users [24]. These systems are designed to deliver accurate and personalized recommendations based on individual health statuses . Considering factors such as a user's health condition, medical history, current medications, and lifestyle habits is essential for ensuring the relevance and effectiveness of healthcare recommendations. By leveraging contextual information effectively, healthcare recommender systems can enhance patient care, promote preventive health measures, and facilitate informed decision-making, ultimately contributing to improved health outcomes and patient satisfaction.
To address the limitations observed in existing healthcare recommender systems, we propose a context-aware healthcare recommendation pipeline that integrates Internet of Things (IoT) sensors data and machine learning techniques for cardiovascular risk prediction. Our approach, referred to as Contextual Wellness Recommender System (CWRS), combines physiological, behavioral, and contextual data to provide personalized, real-time health recommendations.

Table 1 presents a comparative analysis of different healthcare recommender system approaches, highlighting the key advantages and limitations of each method, including our proposed CWRS approach. This comparison emphasizes the distinctive capability of our system to integrate contextual sensing and cardiovascular analytics, thereby enhancing personalization and medical reliability.

As illustrated in Figure 6, the proposed framework is structured as a two-stage pipeline. The first model processes multi-sensor data (mainly accelerometer and gyroscope readings) to recognize the user's physical activity, generating an activity label that captures their behavioral context. This **activity label**, together with three complementary types of information that can affect cardiovascular diseases: **demographic data**, **vital signs** collected from sensors,and **lifestyle factors**.These inputs are then used by a binary classification model to predict the presence or absence of cardiovascular risk.

Table 1. Advantages and Limitations of Healthcare Recommender Systems

Approach	Advantages	Limitations
Collaborative Filtering(CF)	- Personalization based on similar users behavior.	- Suffers from cold-start problem when there's no data for new patients or treatments.

Approach	**Advantages**	**Limitations**
[25]	- Simple to implement with standard algorithms. - Effective for established users with historical data.	- Lacks real-time adaptability to fluctuating conditions. - Limited by profile similarity, not capturing dynamic changes.
Content-Based Filtering(CBRS) [26]	- Recommendations are tailored based on individual patient profiles and specific characteristics, making them personalized and contextually relevant. - Incorporates explicit feedback and known attributes, leading to potentially more accurate recommendations.	- Can suffer from cold start problems when there is limited initial data about a patient or a lack of historical information. - May struggle to adapt to rapidly evolving health conditions or to handle complex cases with insufficient data diversity.
Hybrid Approaches(CF+CBRS) [27]	- Combines collaborative and content-based methods to improve recommendation accuracy. - Leverages multiple data sources, including patient history, to provide more personalized and comprehensive recommendations. - Improves decision-making by integrating various data points and automating parts of the recommendation process.	- Cold start challenges can still occur in cases where there is insufficient data for less common conditions or rare patient profiles. - Incomplete or inconsistent data can negatively impact the system's ability to make accurate recommendations. - Scalability concerns: Managing large and growing datasets can create performance and efficiency challenges over time.
Proposed Approach (CWRS)	- Integrates real-time contextual data for dynamic recommendations. - Can adapt to real-time changes in health status. - Overcomes cold-start issues by incorporating contextual and sensor data. - Useful in continuous patient monitoring and engagement.	- Requires infrastructure for IoT device integration and real-time data collection. - Data quality and consistency may vary depending on sensor accuracy and real-time data transmission. - Potential privacy concerns related to continuous monitoring and real-time health data usage.

7. Experiments

7.1. Model A: User Activity Recognition

Model A is designed to recognize user activities by analyzing motion data collected from a wearable device sensors, such as accelerometers and gyroscopes. The model aims to accurately classify six common daily activities presented in Figure 7, based on multivariate time-series signals. To achieve this, the model leverages a set of features extracted from both the time and frequency domains, capturing essential motion dynamics such as signal magnitude, orientation changes, and periodicity. These features serve as the input to a feed-forward neural network with an output layer of 6 neurons using softmax activation, and trained with the sparse categorical cross-entropy loss function over 10 epochs with a batch size of 64.

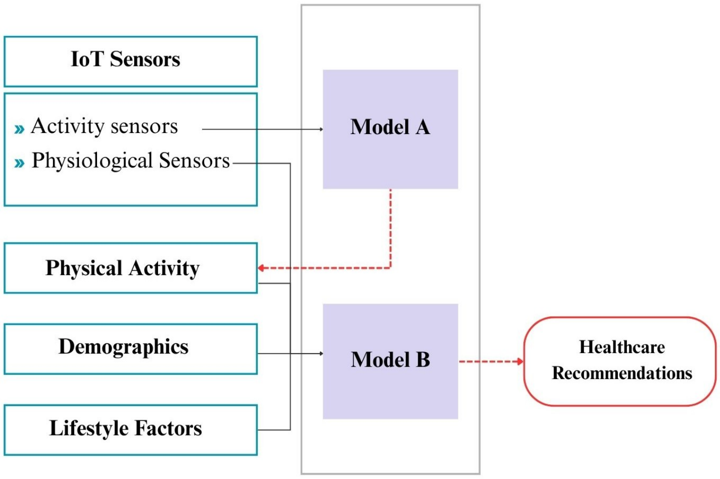

Figure 6. Overview of the proposed framework

This architecture enables efficient learning of motion patterns from multidimensional sensor inputs, allowing the model to achieve high performance in activity classification. The final output corresponds to the activity label with the highest predicted probability, representing the user's most likely current activity.

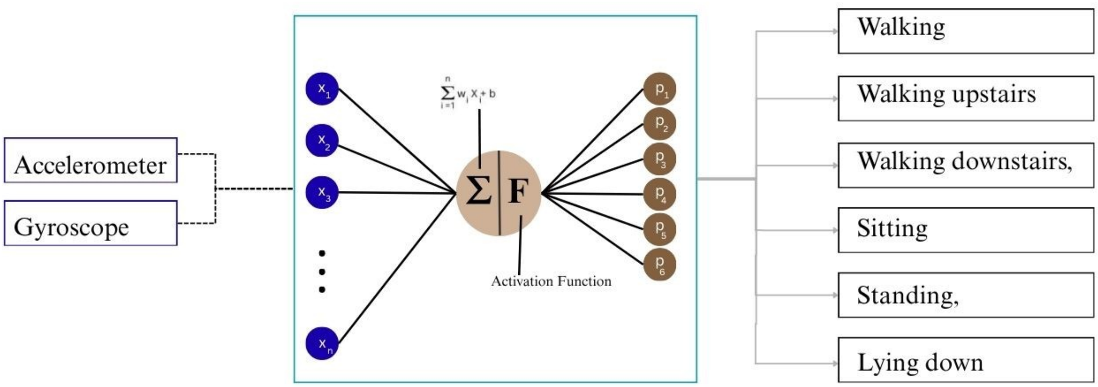

Figure 7. Overview of Model A Architecture

7.2. *Model B: Cardiovascular Risk Prediction and Recommendation Generation*

Model B is responsible for generating personalized health recommendations based on a combination of contextual and clinical features. Specifically, it takes as input the predicted user activity from Model A, along with vital signs, demographic information, and some lifestyle factors. As shown in Figure 8, The goal is to translate these multidimensional indicators into actionable recommendation categories that reflect the user's overall wellness status.

To achieve this, an **XGBoost classifier** was employed to predict the presence or absence of cardiovascular risk. This ensemble tree-based model captures nonlinear relationships between features and it's robust to heterogeneous inputs, while providing strong generalization performance.

Formally, the model learns a mapping:

$$f_B : \mathbb{R}^9 \to \{0, 1\},$$

where $\{0, 1\}$ indicates the predicted cardiovascular risk class (0: low risk, 1: high risk).

The final recommendation $\hat{y}$ is generated using a **rule-based function** applied to the predicted risk class, which translates the binary outcome into interpretable, actionable health advice :

$$\hat{y} = \text{rule(binary prediction)}$$

The recommendations are derived from the World Health Organization (WHO) guidelines on cardiovascular disease prevention. (see Table 2).

Table 2. Personalized Health Recommendations Based on WHO Cardiovascular Prevention Guidelines and Predicted Risk

Condition	**Actionable Recommendation**
Blood Pressure < 70	Check with your doctor if you often feel dizzy or tired.
$70 \leq$ Blood Pressure < 90	Maintain a healthy lifestyle with balanced diet and regular physical activity.
$90 \leq$ Blood Pressure < 110	Watch your salt intake and stay active.
Blood Pressure ≥ 110	Consult your doctor to manage your blood pressure.
100 mg/dL $\leq$ Glucose $\leq$ 125 mg/dL	Choose balanced meals and keep moving daily.
Glucose $>$ 125 mg/dL	Reduce sugar intake and consult a healthcare provider.
200 mg/dL $\leq$ Cholesterol $\leq$ 240 mg/dL	Limit fatty foods and stay active.
Cholesterol $>$ 240 mg/dL	Eat lighter, move more, and get medical advice.
Presence of Cardiovascular Risk	There is a risk of cardiovascular complications! Consult a doctor urgently.

By integrating behavioral, clinical, and demographic information, Model B serves as the decision layer of the overall pipeline, translating fused user data into personalized healthcare recommendations that guide users toward preventive actions and healthier behaviors.

7.3. Datasets

7.3.1. Dataset A : The first dataset is derived from a *Human Activity Recognition* dataset. It contains motion sensor data collected from 30 healthy volunteers aged 19–48 years. While wearing a Samsung Galaxy S II smartphone on their waist, each participant performed six predefined daily activities :

- Walking
- Walking upstairs
- Walking downstairs
- Sitting
- Standing
- Laying

The smartphone's embedded tri-axial accelerometer and gyroscope continuously recorded linear acceleration and angular velocity signals at a constant sampling rate of 50 Hz.

This dataset was made publicly available in a preprocessed form; noise reduction, window segmentation, and signal decomposition were previously applied to the raw accelerometer and gyroscope readings.

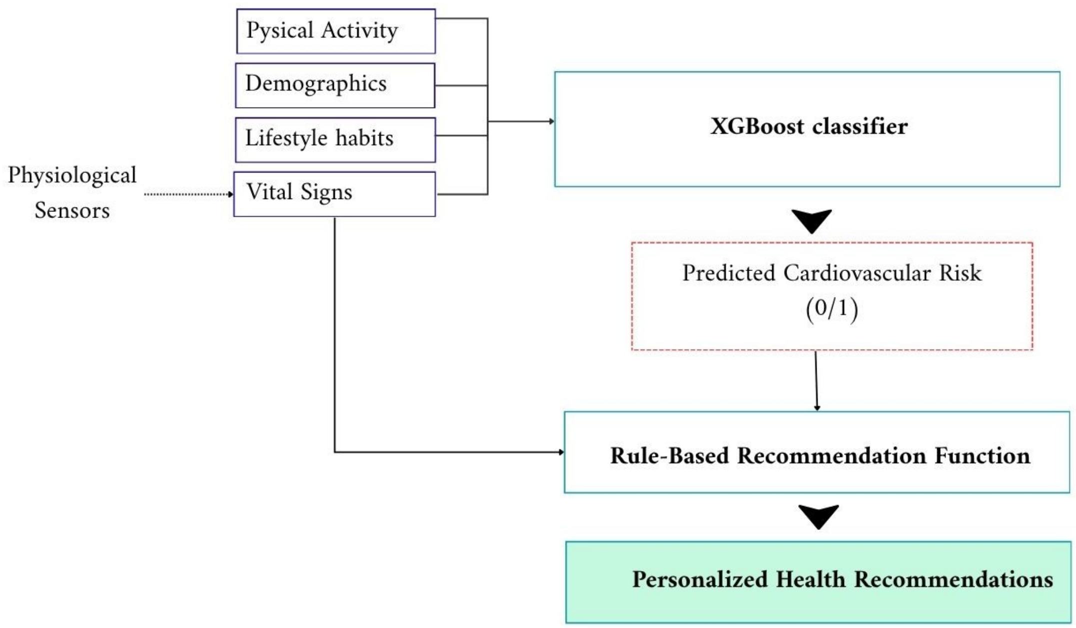

Figure 8. Overview of Model B Architecture

Each record in the final dataset includes a 561 dimensional feature vector extracted from both time and frequency domains, the corresponding activity label, and a subject identifier.

7.3.2. DatasetB : In our recommendation system, the prediction of cardiovascular disease (CVD) serves as the foundation for generating personalized health recommendations for each user. According to a global study [28], cardiovascular diseases (CVDs) are primarily associated with a combination of modifiable and non-modifiable risk factors:

Modifiable Risk Factors:

- Hypertension
- Smoking
- High cholesterol levels
- Diabetes
- Obesity
- Physical inactivity
- Poor diet
- Excessive alcohol consumption
- Stress

Non-modifiable Risk Factors:

- Age
- Sex
- Family history
- Ethnicity

Although certain factors can not be modified, most cardiovascular risk factors can be managed or reduced through lifestyle changes and medical interventions, Healthy eating, maintaining an appropriate weight, regular physical activity, avoiding smoking, and limiting alcohol intake are key preventive measures.

Based on these risk factors, our dataset comprises **70 000 patient records** and includes **11 features** collected during medical examinations, these features were selected to reflect the main cardiovascular risk factors. The features are categorized into **subjective variables** based on user-reported information and **medical measurements**.

To ensure the reliability, consistency, and usability of the dataset for predictive modeling, several preprocessing steps were applied, not only to reduce noise and outliers but also prepare features for meaningful interpretation in the context of medical recommendations:

1. **Outlier Removal**
   Extreme values in continuous variables such as height, weight, systolic blood pressure (ap_hi) and diastolic blood pressure (ap_lo) were removed by excluding the lowest and highest 2.5% of values. This reduces the impact of measurement errors or implausible values on model training.

2. **Feature Derivation: BMI and MAP**
   To enhance the predictive power and stability of the model, two key derived features were computed: the **Body Mass Index (BMI)** and the **Mean Arterial Pressure (MAP)**. Both features summarize important aspects of patient health and allow the model to incorporate clinically meaningful information in a structured way.

   - BMI was calculated from the patient's height and weight using the standard formula:

   $$\text{BMI} = \frac{\text{weight (kg)}}{(\text{height (m)})^2} \tag{1}$$

   BMI values were then discretized into six classes corresponding to standard clinical categories: underweight, normal weight, overweight, obesity class I, obesity class II, and obesity class III. This classification enables the model to capture obesity-related cardiovascular risk in a discrete and interpretable manner.

   - MAP provides a single measure reflecting the average blood pressure of an individual, taking into account both systolic (ap_hi) and diastolic (ap_lo) blood pressures:

   $$\text{MAP} = \frac{2 \times ap\_lo + ap\_hi}{3} \tag{2}$$

   MAP was further categorized into seven classes to represent clinically relevant blood pressure levels, from low to high.
   By transforming continuous vital signs into these derived and categorized features, the dataset becomes more stable and robust for modeling.

3. **Clustering Analysis**
   To capture latent patterns in patient profiles, we applied K-Modes clustering: an unsupervised learning technique suitable for categorical data. This method groups patients with similar characteristics such as BMI class, mean arterial pressure (MAP) class, cholesterol, glucose levels, and lifestyle factors into distinct clusters. Clustering was performed separately for female and male subgroups to account for sex-specific patterns. Each patient was assigned a cluster label, which can be used as an additional feature in predictive modeling.
   The importance of these cluster labels and other features is illustrated in Figure 9 which shows the top 10 features ranked by the XGBoost model. Notably, the cluster labels exhibit a very significant influence, followed by MAP class, BMI class. This demonstrates that the preprocessing and feature engineering steps, including clustering and the creation of medically relevant classes, improve the model's ability to recognize patterns in patient data and enhance cardiovascular risk predictions.

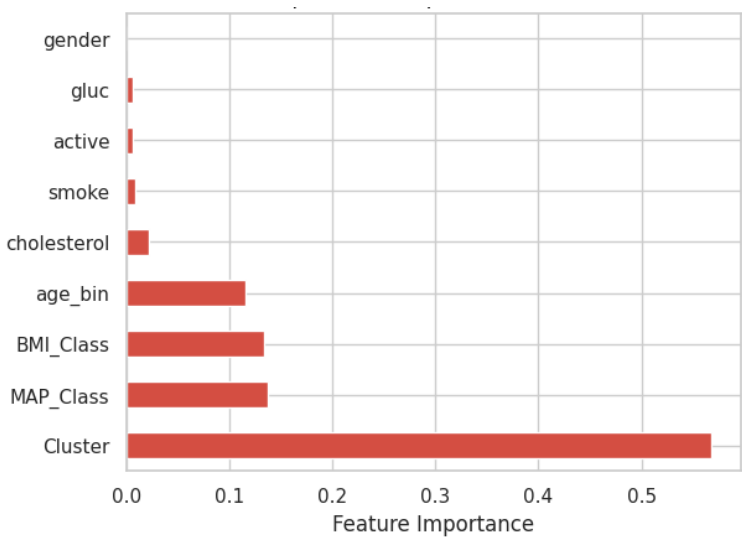

Figure 9. Feature Importance Ranking from the XGBoost Model for Cardiovascular Risk Prediction

7.4. Experimental Results

7.4.1. Model A: The physical activity Prediction MLP was trained over 10 epochs, achieving a final training accuracy of 95.26% and a validation accuracy of 92.81%, with corresponding loss values of 0.13 and 0.18, respectively. The learning curves in Figures 10 and 11 illustrates a steady improvement in both training and validation performance, with convergence observed after approximately the sixth epoch. This indicates effective generalization and the absence of significant overfitting. Such a performance level demonstrates the model's strong ability to accurately distinguish between different activity classes based on inertial sensor data (accelerometer and gyroscope signals). This result is of particular importance since the predicted activity class serves as a key input to the second-stage model responsible for cardiovascular risk prediction and personalized health recommendation generation. Consequently, achieving high accuracy in this stage ensures that subsequent risk assessments and recommendations are grounded in reliable behavioral context information, enhancing the overall precision and trustworthiness of the proposed IoT-based healthcare recommendation pipeline.

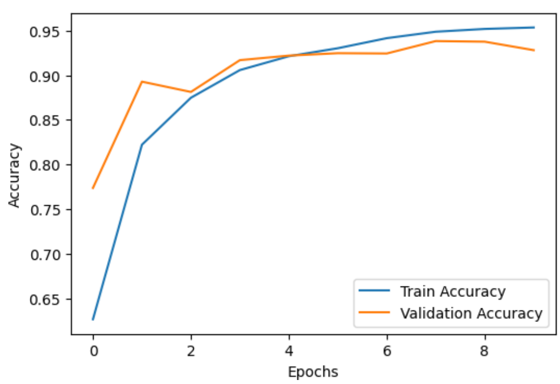

Figure 10. Model_A Accuracy Curve

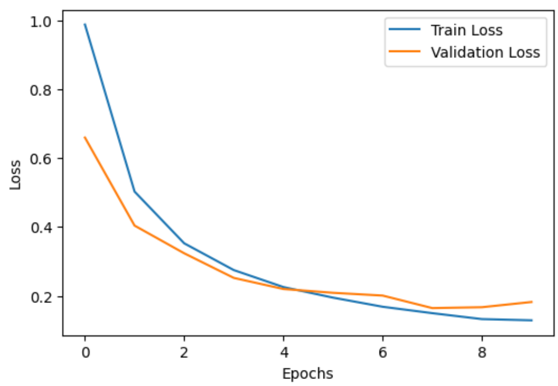

Figure 11. Model_A Loss Curve

7.4.2. Model B: To evaluate the performance of our cardiovascular risk prediction model, we trained and tested multiple classification algorithms, including **Random Forest, Logistic Regression, Support Vector Machine,**

K-Nearest Neighbors, and **XGBoost**. Each model was assessed using several evaluation metrics such as accuracy, precision, recall, and F1-score to ensure a comprehensive understanding of their performance. Table 3 presents a comparative evaluation of multiple classifiers applied to the cardiovascular risk prediction task.

Table 3. Performance comparison of different classifiers for cardiovascular risk prediction

Model	Accuracy	Precision	Recall	F1-score
Random Forest	0.869	0.87	0.87	0.87
Logistic Regression	0.827	0.835	0.825	0.83
Support Vector Machine	0.837	0.84	0.835	0.82
K-Nearest Neighbors	0.852	0.855	0.855	0.85
XGBoost	0.874	0.875	0.87	0.87

Overall, all models achieve relatively high performance, with accuracy values ranging from 82.7% to 87.4%. Among the tested algorithms, XGBoost demonstrates the best performance across most metrics, achieving an accuracy of 87.4%, precision of 0.875, recall of 0.87, and F1-score of 0.87. Random Forest and K-Nearest Neighbors also perform competitively, while Logistic Regression and Support Vector Machine show slightly lower scores. To further assess and visualize the discriminatory power of each model, the ROC curves were plotted 12, highlighting that XGBoost not only achieves high overall metrics but also exhibits the highest area under the curve (AUC), confirming its superior ability to distinguish between positive and negative cardiovascular risk cases.The close values of precision, recall, and F1-score 3 indicate that the models maintain a balanced performance between correctly identifying positive and negative cases, minimizing the risk of bias toward either class,This proximity can be attributed primarily to the quality and structure of the dataset, which is well-balanced, preprocessed, and contains informative features. These results suggest that ensemble-based methods like XGBoost and Random Forest are particularly effective for this cardiovascular risk classification task.

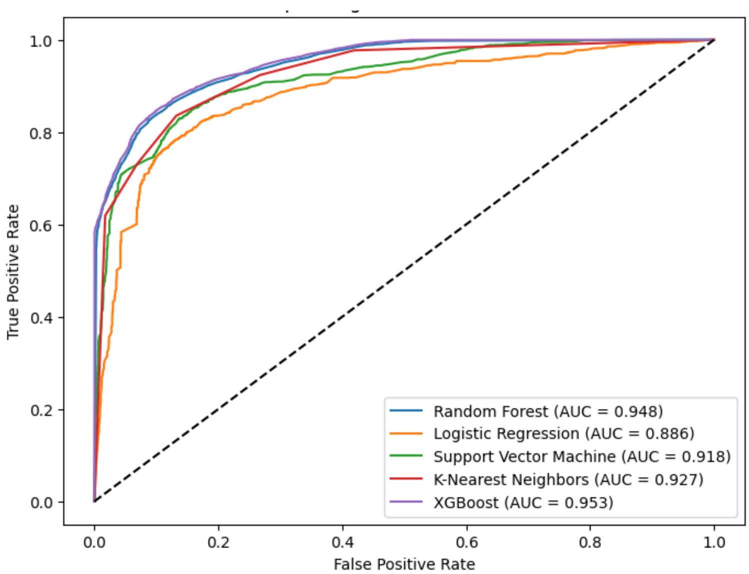

Figure 12. Receiver Operating Characteristic (ROC) Curves for Cardiovascular Risk Prediction Models

Given its superior performance across multiple evaluation metrics, the **XGBoost** algorithm was selected as the final model for cardiovascular risk prediction. To ensure its robustness, several evaluation procedures were applied; the model achieved an accuracy of 87.39% on the test set, with the optimal number of estimators found via GridSearchCV being 100. A 95% confidence interval for accuracy, estimated through bootstrap resampling, ranged from 86.90% to 87.88%, indicating stable generalization performance.

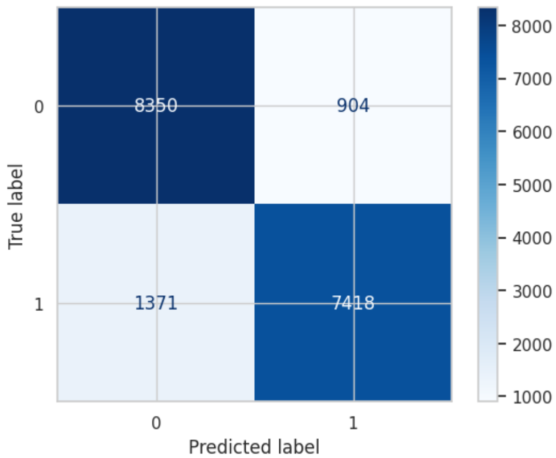

Figure 13. Confusion Matrix of the XGBoost Cardiovascular Risk Prediction Model

The confusion matrix 13 shows that the model correctly classified a majority of both positive and negative cases, with 7418 true positives and 8350 true negatives, while misclassifications remained relatively low (904 false negatives, 1371 false positives).

We also evaluated the calibration curve 14, which demonstrated excellent alignment with the ideal calibration line, confirming that the predicted probabilities closely reflect the true outcome frequencies. Furthermore, the learning curve 15 indicated no signs of overfitting, suggesting that the model generalizes well to unseen data.

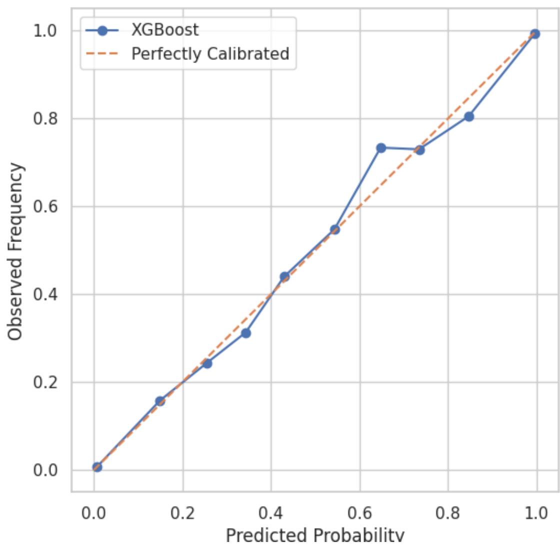

Figure 14. Calibration Curve of the XGBoost Cardiovascular Risk Prediction Model

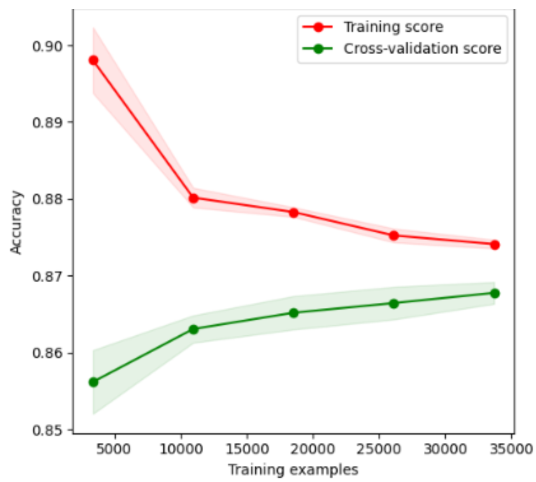

Figure 15. Learning Curve of the XGBoost Cardiovascular Risk Prediction Model

The obtained results from the cardiovascular risk prediction model highlight the robustness and reliability of the proposed system. The high accuracy and balanced precision–recall values demonstrate that the model effectively differentiates between individuals at high and low cardiovascular risk based on features derived from physiological and behavioral indicators.

From a medical perspective, these results suggest that the model successfully captures the underlying correlations between activity-related metrics, physiological patterns, and cardiovascular risk factors, also the model's calibration performance further reinforces its reliability, ensuring that predicted probabilities align closely with real-world event likelihoods (an essential property for clinical decision support systems).

8. Discussion

In this study, we developed a two-stage, context-aware healthcare recommender system aimed specifically at cardiovascular risk prediction and personalized health guidance. The first stage of our pipeline focuses on recognizing user's physical activity using data captured from sensors (accelerometer and gyroscope), the proposed activity classification model (Model A) achieved a high performance, demonstrating its strong ability to distinguish between different physical activities. This robust performance is crucial, as the predicted activity class serves as a contextual input to the second-stage model.

The second stage combines the output of the activity recognition model with additional patient data, including demographics, vital signs (e.g., blood pressure, cholesterol, glucose), and lifestyle factors such as alcohol and smoking habits, to predict cardiovascular risk via a binary classification model. Our evaluation involved several machine learning algorithms(Random Forest, Logistic Regression, Support Vector Machine, K-Nearest Neighbors, and XGBoost) tested on a large dataset with several preprocessing steps. The results show that the high-quality dataset explains why even simpler models achieved relatively close performance, underscoring the importance of data preparation in medical predictive models.

For the best-performing model (XGBoost), we conducted additional analyses to assess its overall reliability and consistency. The evaluation confirmed that the model demonstrates stable predictive behavior, accurate probability estimation, and strong generalization performance, making it suitable for reliable cardiovascular risk assessment.

From a medical perspective, this approach has several important implications, by integrating continuous activity monitoring with vital signs and lifestyle factors, our system provides a more comprehensive, individualized assessment of cardiovascular risk. The rule-based recommendation layer translates model predictions into actionable advice, enabling users to take preventive measures such as dietary adjustments, regular physical activity, and monitoring of blood pressure and lipid levels. This combination of predictive analytics and personalized guidance can improve early detection of cardiovascular risk factors, enhance patient engagement in healthy behaviors, and ultimately contribute to reducing the incidence of cardiovascular diseases.

In summary, our work highlights the importance of combining multi-source contextual data with machine learning models in a medical context. Beyond prediction accuracy, we emphasize interpretability and actionable recommendations, which are critical for translating computational outputs into meaningful healthcare interventions.

9. Conclusion

This study proposed an integrated IoT–based healthcare recommendation framework that leverages contextual and clinical information to deliver personalized health insights. The approach combines motion data from wearable sensors with physiological and demographic indicators to generate meaningful, user–specific recommendations aimed at supporting preventive healthcare and cardiovascular well–being.

By connecting behavioral context with health status indicators, the proposed framework contributes to the advancement of context–aware healthcare Recommender Systems. It highlights the potential of IoT and AI technologies to enable continuous monitoring, early detection, and adaptive health guidance.

For future work, we plan to expand the contextual scope of our recommendation mechanism by integrating additional personal and situational factors such as gender, fasting state, post–meal periods, and daily routine context. This enhancement aims not only to enable finer–grained personalization but also to ensure that the generated recommendations are more medically accurate and aligned with each user's physiological and lifestyle conditions. In addition, we aim to explore the integration of data from multiple IoT sources, including smartphones and wearable devices, using advanced sensor fusion techniques. This direction is expected to further enhance the robustness and precision of contextual recommendations.

Furthermore, future experiments will involve testing the framework on real–world IoT sensor streams and integrating it into a mobile application for real–time, interactive health monitoring.

REFERENCES

1. G. Adomavicius and A. Tuzhilin, *Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions*, IEEE Transactions on Knowledge and Data Engineering, vol. 17, pp. 734–749, July 2005. DOI: 10.1109/TKDE.2005.99.
2. G. Adomavicius, R. Sankaranarayanan, S. Sen, and A. Tuzhilin, *Incorporating contextual information in recommender systems using a multidimensional approach*, ACM Transactions on Information Systems (TOIS), vol. 23, no. 1, pp. 103-145, 2005.
3. A. Yadav and S. Shukla, *Context-aware recommender systems: A review of the state of the art techniques*, Journal of Intelligent Information Systems, vol. 53, no. 2, pp. 195-223, 2019.
4. F. Ricci, L. Rokach, and B. Shapira, *Introduction to Recommender Systems Handbook*, Springer, 2011.
5. C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos, *Context-aware computing for the Internet of Things: A survey*, IEEE Communications Surveys & Tutorials, vol. 16, no. 1, pp. 414-454, 2014.
6. R. Burke, A. Felfernig, and M. Göker, *Recommender Systems: An Overview*, AI Magazine, vol. 32, pp. 13–18, Sept. 2011.
7. P. Singh, P. Dutta Pramanik, A. Dey, and P. Choudhury, *Recommender Systems: An Overview, Research Trends, and Future Directions*, International Journal of Business and Systems Research, vol. 15, pp. 14–52, 2021.
8. F. Mansur, V. Patel, and M. Patel, *A review on recommender systems*, in Proc. of the IEEE International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), pp. 1–6, Mar. 2017.
9. J. Smith, *Understanding Netflix's recommendation algorithm*, Journal of Digital Media, vol. 14, no. 2, pp. 56-72, 2017.
10. P. Lops, D. Ferrante, and M. de Gemmis, *Content-based recommender systems: From the user model to the recommendation*, In *Recommender Systems Handbook*, pp. 73-105. Springer, 2011.
11. M. J. Pazzani and D. Billsus, *Content-based recommendation systems*, In *The adaptive web*, pp. 325-341. Springer, 2007.
12. N. Lathia, S. Hailes, and L. Capra, *The role of social networks in the cold start problem*, ACM Transactions on Intelligent Systems and Technology, vol. 4, no. 3, pp. 1-20, 2013.
13. Y. Koren, *Factorization meets the neighborhood: a multifaceted collaborative filtering model*, Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 426-434, 2008.
14. Y. Zhang and L. Chen, *Deep learning for recommender systems: A survey and new perspectives*, ACM Computing Surveys (CSUR), vol. 52, no. 1, pp. 1-38, 2019.
15. T. H. Kuo and Y. C. Lee, "Recommender systems for healthcare: A systematic review," *Expert Systems with Applications*, vol. 94, pp. 123-139, 2018.
16. Y. Aharony and J. Tzeng, "The impact of wearable technology on health behavior: A systematic review," *International Journal of Medical Informatics*, vol. 132, p. 103972, 2019.
17. S. Shakya and R. Tiwari, "Wearable technologies for health and wellness: A review of the state-of-the-art," *Journal of Biomedical Informatics*, vol. 92, p. 103134, 2019.
18. G. Adomavicius, B. Mobasher, F. Ricci, and A. Tuzhilin, *Context-Aware Recommender Systems*, AI Magazine, vol. 32, pp. 67–80, 2011.
19. L. Tamine-Lechani and S. Calabretto, *Recherche d'information contextuelle et web*, 2008.
20. S. Kulkarni and S. Rodd, *Context Aware Recommendation Systems: A review of the state of the art techniques*, Computer Science Review, vol. 37, p. 100255, 2020. DOI: 10.1016/j.cosrev.2020.100255.
21. D. Nawara and R. Kashef, *Context-Aware Recommendation Systems in the IoT Environment (IoT-CARS) – A Comprehensive Overview*, IEEE Access, vol. 9, pp. 144270–144284, 2021. DOI: 10.1109/ACCESS.2021.3122098.
22. O. Stitini, F. Ouakasse, S. Rakrak, S. Kaloun, and O. Bencharef, *Combining IoMT and XAI for Enhanced Triage Optimization: An MQTT Broker Approach with Contextual Recommendations for Improved Patient Priority Management in Healthcare*, International Journal of Online and Biomedical Engineering (iJOE), vol. 20, pp. 145–162, 2024. DOI: 10.3991/ijoe.v20i07.47483.
23. F. Ouakasse, O. Stitini, and S. Rakrak, *A Context-Aware Framework to Manage the Priority of Injured Persons Arriving at Emergencies*, International Journal of Online and Biomedical Engineering (iJOE), vol. 19, no. 08, pp. 144–162, 2023. DOI: 10.3991/ijoe.v19i08.3919.
24. O. Stitini, S. Kaloun, and O. Bencharef, *Latest Trends in Recommender Systems applied in the medical domain: A Systematic Review*, 2020.
25. J. Medina, O. Apolinario-Arzube, H. Luna Aveiga, K. Lagos-Ortiz, M. Paredes-Valverde, and R. Valencia-García, "A Collaborative Filtering Based Recommender System for Disease Self-management," in *Proceedings of the 7th International Conference on Information Technology and Intelligent Transportation Systems*, Springer, pp. 60–71, 2017. DOI: 10.1007/978-3-319-67283-0_5.
26. S. Rana, M. Dey, J. Prieto, and S. Dudley, *Content-Based Health Recommender Systems*, in *Advances in Ubiquitous Computing: Future Paradigms and Directions*, Chapter 11, pp. 215–236, Wiley, 2020. DOI: 10.1002/9781119711582.ch11.
27. Q. Zhang, G. Zhang, J. Lu, and D. Wu, *A Framework of Hybrid Recommender System for Personalized Clinical Prescription*, in *Proceedings of the 2015 IEEE International Conference on Intelligent Systems and Knowledge Engineering (ISKE)*, pp. 189–195, 2015. DOI: 10.1109/ISKE.2015.98.
28. J. Nevins et al., *Global Effect of Modifiable Risk Factors on Cardiovascular Disease and Mortality*, New England Journal of Medicine, vol. 388, no. 5, pp. 401–412, 2023. [Online]. Available: https://www.nejm.org/doi/full/10.1056/NEJMoa2206916