SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V11I11P122

Volume 11 Issue 11, 222-230, November 2024

Review Article

Machine Learning and Smart Devices for Prediction of Heart Disease, Diabetes, and Obesity: Systematic Review

Kirti Gupta[1*], Pardeep Kumar[2], Shuchita Upadhyaya[3], Shalini Aggarwal[4]

[1,2,3]*Department of Computer Science & Applications, Kurukshetra University, Haryana, India.*
[4]*Department of Computer Science, S.U.S. Government College, Karnal, Haryana, India.*

**Corresponding Author : kirtigupta@kuk.ac.in*

Received: 12 September 2024 Revised: 13 October 2024 Accepted: 11 November 2024 Published: 30 November 2024

Abstract - *Heart disease, diabetes, and obesity are major global health issues and are increasing day by day. There is a need for effective predictive strategies. Machine learning is vital for early detection and diagnosis. It provides timely and individualized healthcare interventions. This analysis evaluates optimal ML algorithms for heart disease, diabetes, and obesity. It identifies key risk factors and highlights the role of AI in healthcare. The study follows PRISMA guidelines. A broad search using Google Scholar and PubMed identified 30 studies from 2014 to 2024. The article explores the symptoms and aftereffects of diabetes, obesity, and heart disease. It reviews effective ML methods for forecasting these issues. AI-based methods help medical professionals diagnose diseases. Research shows that AI tools improve diagnosis speed and accuracy. This leads to more personalized treatment and better outcomes.*

Keywords - *Diabetes, Healthcare, Heart disease, Machine Learning (ML), Obesity.*

1. Introduction

Cardiovascular disease, along with diabetes and obesity, are prominent health issues worldwide [1]. These conditions have a negative impact not only on healthcare systems but also on global economies [2]. It is widely acknowledged that timely identification and intervention are crucial in enhancing patient prognosis [3]. Taking steps can mitigate the lasting repercussions of these illnesses.

Machine Learning (ML), a subset of intelligence, is a tool that plays a significant role in predicting and detecting a range of illnesses like heart disease, diabetes mellitus, and obesity [4]. ML models can uncover trends within datasets to assess the likelihood of these ailments [1]. This study broadens our understanding to aid in the detection of risks.

The use of devices and smartphones is increasingly common nowadays as they offer access to a wealth of health-related information [3]. ML algorithms can predict and monitor various health conditions by collecting data from these gadgets, like blood pressure readings and glucose levels and monitoring activity levels and sleep patterns [4]. This study provides an in-depth overview of the status of machine learning applications in predicting obesity, diabetes, and heart disease. We will explore the ML algorithms in disease prediction tasks and the potential benefits and challenges associated with predicting and managing these illnesses.

2. Related Research Works

2.1. A Critical Review of Heart Disease Prediction

Despite the significant progress in medical science, heart disease is still the leading cause of death worldwide. This underscores the pressing need for adequate early detection and intervention. Multiple studies have examined the application of ML in improving both diagnostic accuracy and its identification. Therefore, ML is a very prominent approach for predicting cardiac disease. Various ML algorithms are used in the literature to predict heart disease. This includes advanced techniques such as Neural Networks, Deep Learning (DL), and more traditional ones like SVM and Logistic Regression.

Researchers have used different data types, including health records and demographic information. In [4], the authors incorporate lifestyle data and ECG information as important predictors. Although [5] shows ML can efficiently predict heart disease. However, the best methods and data sources must be evaluated according to the study issue, target group, and resources.

Some models are only useful in the case of transfer-learning techniques. Due to differences in risk factors, demographics, data collection methods, and markers, we cannot realize model potential properly. Moreover, various ML algorithms are so "high-fidelity" that we cannot understand how this or that model works.

This is important because if the future can be predicted, it might help us gain insight into the principles behind predictions. Therefore, patients are often categorized according to their propensity for successful outcomes, and risk assessments guide treatment decisions. This lack of transparency has created many problems for healthcare providers using these models.

Biases can influence predictions in the data used for training. The under-representation of particular groups or biases in healthcare systems may exacerbate health disparities. Robust bias assessment is required to ensure fair heart disease prediction models. Make sure to have processing strategies that pivot toward equity, privacy, and informed consent to trust in the deployment of AI. To ameliorate these, it is necessary to cultivate confidence and ethics in ML.

2.2. A Critical Review of Diabetes Prediction

There are several promising outcomes of Machine learning for diabetes prediction. These are in response to the increasing incidence of diabetes and the benefit of early detection [4, 6]. To build better ML models for the prediction of diabetes, good knowledge and understanding of work done before pushing the field of study [3]. This can be used to profile strengths, weaknesses, and possible future improvement routes.

Diabetes prediction uses a collection of machine learning algorithms. It comprises conventional methods such as SVM and Logistic Regression. It also includes more sophisticated methods such as Neural Networks and Deep Learning [7, 8]. However, [9-11] show that selecting the perfect algorithms for precise classification relies highly on the particular dataset and features employed besides their target population.

Data heterogeneity across research is a fundamental problem in diabetes prediction [6]. Variations in data collection, risk variables, and demographics may restrict model generalizability. It is possible that a model developed for one ethnic group will not work effectively for another. The necessity of including various risk factors and demographics in prediction models is emphasized in [12].

More relevant features are crucial in joining accurate and interpretable models [13]. For instance, BMI, family history, and age are important traditional risk factors in literature surveys. Some surveys have also involved more sophisticated features from electronic health records or wearable sensors [14]. However, providing more features than are required for a model will likely lead to overfitting the model and make it less interpretable. The key challenge is to find the balance between accuracy and interpretability.

Diabetes Prediction Model generalization for clinics remains an important issue. This problem is the same as that in the prediction of heart disease. Models trained on only one dataset may not scale well to other settings because data collection methods, demographics, and risks may differ in different datasets. Moreover, certain advanced ML algorithms operate as a black box, and these models may compromise interpretability by not seeing some important predictions. A lack of transparency could create a hurdle for clinical uptake clinicians generally should be able to interpret risk assessments that inform their treatment decisions.

All AI healthcare applications remain uncertain regarding ethical and potential bias issues [6]. That is, using training data biases increases the variance in prediction accuracy. Fairness, transparency, and accountability are important criteria in developing and deploying diabetes prediction models.

2.3. A Critical Review of Obesity Prediction

As the global burden of obesity increases, the enthusiasm for prediction and personal intervention from machine learning has also increased. Motivated by these promising developments, this article aims to present a comprehensive evaluation of the study and development of ML in prediction.

Numerous ML model algorithms are used to predict obesity, such as diabetes and heart disease detection [15]. From traditional regression models to more advanced deep learning architectures, these algorithms span a wide spectrum [2]. Researchers have explored several parameters, such as living patterns, dietary habits, natural markers, and environmental factors [16]. This points to the multi-factorial nature of obesity.

In common with all health prediction tasks, resilience to the generalizability problem is likely a factor for tasks such as obesity. Data collection methods, culture-affecting lifestyle choices, and genetic profiles that predetermine the response towards certain changes are all factors leading to models trained on specific datasets with limited applicability [17].

Complex ML algorithms can also be more difficult to interpret since they are not always as transparent. They make it difficult to decipher what is being predicted and thus reduce clinical trust in and acceptance of the methodology.

Unfortunately, if the representation of specific ethnic groups or populations with different socioeconomic backgrounds in the training data is poor, biases could raise concern where it may deepen existing health disparities and make predictions less accurate for certain populations.

For example, [17] points out that in research on obesity, there is currently very limited code sharing, which can prevent reproducibility and increase bias. This includes controlling for biases and augmenting data in obesity prediction models, as well as creating fairness-aware algorithms.

Though accurate prediction is an important first step, the ultimate goal of using ML in obesity research is to identify effective personalized interventions. This entails going beyond just risk scores and support for what the best action might be. [16] highlights the promise of ML in uncovering geographically proximate risks and subsequent interventions.

2.4. Combining Predictive Models for Multiple Conditions

Given the shared risk factors and relationships between CVD, diabetes, and obesity, developing integrated predictive models would be beneficial. These conditions arise from genetics, lifestyle, and environmental factors. The co-occurrence of these conditions has also been the subject of study with different applications for ML prediction work reported [18-20].

Author [19] presented a majority voting method that merges machine learning techniques to forecast and categorize obesity effectively by utilizing demographic information and clinical data in conjunction with genetic markers to enhance the prediction accuracy for the simultaneous occurrence of heart disease and diabetes conditions while showing the significance of feature engineering and model explanation and highlighting the capabilities of deep learning models, for risk prediction related to heart disease, obesity, and diabetes.

3. Research Methods

3.1. Review Protocol

This paper helps to identify the most effective ML algorithm used for heart disease, diabetes, and obesity prediction that captured data from smart devices. Identification and diagnosis of these diseases as early as possible is essential.

Therefore, ML algorithms are an appropriate solution to predictions of these diseases. This systematic review applies PRISMA guidelines 2020, as shown in Figure 1. A comprehensive search uses Google Scholar and PubMed. 30 articles published between January 2014 and August 2024 are included to address these diseases.

3.2. Identifying Research Questions

This study explores the following study topics.

- What are the key symptoms and aftereffects of Heart Disease, Diabetes, and Obesity?
- What are the most effective ML methods for forecasting Heart Disease, Diabetes, and Obesity?
- How do intelligent systems help doctors diagnose diseases?

3.3. Inclusion & Exclusion Criteria

The PRISMA 2020 guidelines filter our search results systematically, ultimately identifying 30 scientific articles suitable for inclusion in our review study, as illustrated in Figure 1. Our initial search yielded 2,846 articles from the PubMed database and 582 from Google Scholar. After applying an additional filter for "Free Full Text", the number reduces to 2,196. We then refined the results by selecting only "Systematic Review" articles published within the last 10 years, narrowing the list further to 54 papers. After screening the titles and abstracts to exclude unrelated studies, we arrived at 30 relevant scientific articles for our review.

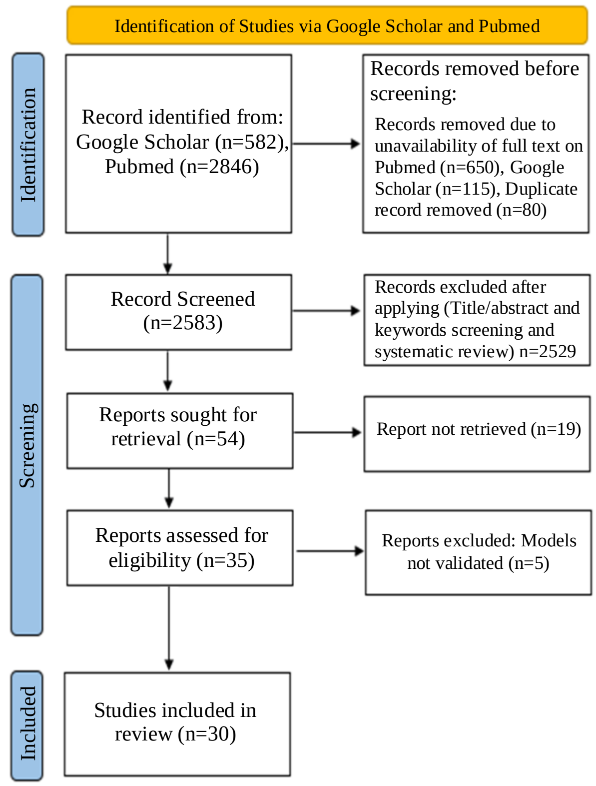

Fig. 1 The PRISMA flow diagram

The Google Search query used in this study is:

Allintitle: "Machine learning" AND ("heart disease" OR diabetes OR obesity AND "Prediction")

Custom Range=2014-2024

The Boolean search string for the PubMed Database used in this study is:

(Prediction[Title/Abstract] OR Factors[Title/Abstract] OR Parameters[Title/Abstract]) AND (Heart*[Title/Abstract] OR diabetes*[Title/Abstract] OR Obesity[Title/Abstract]) AND (machine learning[Title/Abstract]) AND (medical[Title/Abstract] OR biomedical[Title/Abstract] OR health[Title/Abstract] OR healthcare[Title/Abstract] OR clinical[Title/Abstract])

3.4. Search Strategy

This investigation centered on pinpointing themes and patterns in the domain.

- We investigated machine learning algorithms used to predict heart disease, diabetes, and obesity, ranging from statistical methods to advanced deep learning techniques.
- We looked into strategies to detect and address biases in models to ensure fairness and equality when evaluating risks.
- We delved into utilising machine learning models to create insights for interventions and enhance disease management strategies.

Author [2] gives a comprehensive account of ML approaches to obesity prediction, and [16] discusses the importance of environmental factors in conceptualizing obesity risk geographically. Author [17] recommends code sharing and transparency to account for potential biases in obesity research. Through a thorough literature review, this article will help identify key challenges and opportunities in previous research for guiding future endeavors toward more effective and equitable solutions to addressing this global health issue. An overview of some reviewed studies is shown in Table 1.

The application of ML is expected to help physicians interpret patient data and implement optimal algorithms for their datasets. Additionally, ML can generate evidence that helps healthcare providers manage individual situations requiring invasive procedures such as catheterization.

The variety of the identified DL models in the smart devices confirms that, for the selection of the most suitable DL Architecture, the one-size-fits-all approach does not apply. The findings have also been supported by reviews [23, 24]. The models used in the smart device studies and the methodologies employed did not provide information regarding the consistency of the statements.

Deep learning models have demonstrated remarkable potential in predicting these diseases, especially in medical imaging, electronic health records, genomics, and bio-signal analysis [21-23]. Convolutional neural networks excel in processing medical images, while recurrent neural networks and long short-term memory models handle time-series data to predict disease progression. Autoencoders and generative adversarial networks serve different purposes in the field of ML, with autoencoders focusing on feature extraction and data augmentation.

Nevertheless, obstacles such as limited data availability, interpretability of models, and generalization to new scenarios persist. Recent breakthroughs, such as transfer learning, explainable artificial intelligence, and multi-modal learning, are crucial in overcoming these limitations. Looking ahead, personalized medicine and federated learning are the future directions, making deep learning a promising tool in healthcare. However, more research is required to integrate it fully into clinical practice.

4. Research Question Results

Heart disease, diabetes, and Obesity are intricately connected, as shown in Figure 2, forming a harmful cycle. Excessive body fat, especially around the abdomen, causes insulin resistance, which significantly raises the risk of developing type 2 diabetes. In turn, this illness can harm blood vessels and nerves, raising the risk of heart disease, which is frequently made worse by other diabetes-related problems like elevated blood pressure and abnormal cholesterol levels. Moreover, obesity raises blood pressure, cholesterol, and inflammatory markers, which puts greater strain on the heart and causes diseases like atherosclerosis. Because of the way these disorders are tied to one another, managing and preventing these related health issues requires addressing lifestyle variables holistically. Tables 2, 3, and 4 provide a concise summary of heart disease, diabetes, and obesity, three interrelated illnesses influenced by many causes.

Table 2. Factors affecting diabetes

Factor	Descriptions
Genetics	A family history of diabetes can influence the likelihood of developing the condition.
Diet	Consuming high amounts of sugars and refined carbs can raise the risk of diabetes.
Physical Activity	Engaging in regular exercise supports insulin sensitivity and helps regulate blood sugar levels.
Obesity	Excessive body weight, especially around the abdomen, significantly increases the risk of type 2 diabetes.
Age	Diabetes probability rises with age, typically after 45.
Insulin Resistance	In type 2 diabetes, cells become less responsive to insulin, resulting in higher blood sugar levels.
Gestational Diabetes	Mainly occurs in women during pregnancy. They have a greater probability of developing type 2 diabetes occurrence later on.
Sleep Patterns	Poor sleep quality and sleep disorders can disrupt glucose metabolism and reduce insulin sensitivity.

Table 3. Factors affecting heart disease

Factor	Descriptions
Genetics	A family history of heart disease can heighten the risk of developing it.
Diet	Consuming large amounts of saturated fats, trans fats, and cholesterol can lead to heart disease.
Physical Activity	A lack of physical activity can heighten the risk.
Smoking	Using tobacco can harm blood vessels and increase the likelihood of heart disease.
Alcohol Consumption	Drinking excessively can raise blood pressure and contribute to heart disease.
Blood Pressure	Elevated BP places additional stress on the heart and arteries.
Cholesterol Levels	low levels of HDL (good cholesterol) and High levels of LDL (bad cholesterol)
Diabetes	High blood sugar levels associated with diabetes can damage blood vessels, raising the risk of heart disease.
Stress	Ongoing stress can contribute to heart disease through factors like poor eating habits and smoking.

Table 4. Factors affecting obesity

Factor	Descriptions
Diet	Consuming high calories, particularly from sugary and fatty foods, can lead to weight gain.
Physical Activity	A lack of physical activity is a major contributor to obesity.
Genetics	Genetic traits can impact how the body accumulates and processes fat.
Hormonal Changes	Hormones like leptin and ghrelin play key roles in regulating hunger and metabolism.
Sleep	Inadequate sleep can disrupt hormones that control appetite, resulting in weight gain.
Stress	Long-term stress can lead to emotional eating and weight gain.
Medications	Some medications can cause an increase in weight.
Socioeconomic Factors	Limited financial resources may restrict access to healthy food options and opportunities for exercise.

Interconnections:

- Heart Disease and Obesity: Excess weight strains the heart and increases the risk of high blood pressure, high cholesterol, and diabetes.
- Diabetes and Heart Disease: Diabetes accelerates the development of atherosclerosis (hardening of the arteries) and other heart-related issues.
- Obesity and Diabetes: Excess fat, especially visceral fat, contributes to insulin resistance, leading to type 2 diabetes.

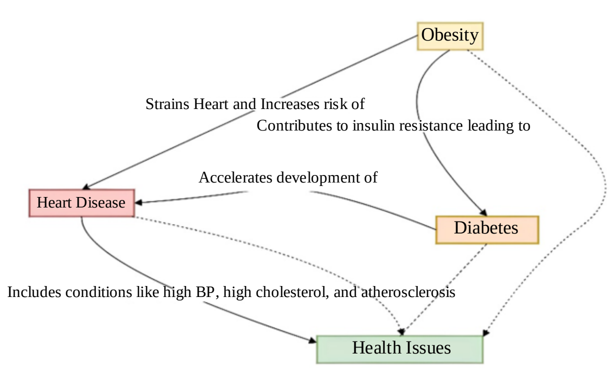

Fig. 2 Interconnections between heart disease, diabetes, and obesity

Addressing these factors through lifestyle changes, such as a healthy diet, increasing physical tasks, avoiding stress, and getting regular medical check-ups, can help manage and reduce the risk of these conditions.

Intelligent systems, especially those that use machine learning and Artificial Intelligence (AI), greatly improve physicians' ability to diagnose diseases by offering several important benefits. These computers can process significantly more medical data than a human could reasonably assess in a reasonable period. This data includes patient histories, test findings, imaging, and genetic information. Intelligent systems enhance diagnostic speed and accuracy, enabling earlier and more accurate disease identification by seeing patterns and correlations that the human eye might miss. AI algorithms are used in imaging technologies to identify abnormalities in scans, frequently resulting in the early diagnosis of diseases like cancer.

Moreover, by continuously learning from fresh data and incorporating evidence-based information into clinical procedures, these systems assist physicians in staying current with the most recent medical research and guidelines. Furthermore, decision support systems can provide alternative diagnoses and courses of therapy, helping doctors make well-informed choices specific to each patient. Intelligent systems supplement physicians, boosting patient care and increasing diagnostic results through sophisticated data analysis and decision-making assistance.

5. Conclusion and Future Directions

This detailed review delves into using ML algorithms and intelligent gadgets to predict and handle heart disease, diabetes, and obesity. It covers aspects that impact these health issues in ML and discusses the role of AI-driven methods in diagnosing diseases.

The analysis points out that heart disease and obesity are linked by causes like genetics and lifestyle choices while also considering factors as crucial influencers. For heart condition prediction ensemble techniques, neural networks and SVM are widely used ML algorithms to study intricate data sets, identify high-risk individuals, and forecast disease onset. These models can tackle incomplete data, varying data types, and merging clinical insights for improved understanding and practical medical use. For diabetes analysis, in healthcare settings like clinics and hospitals, different models such as regression, decision trees, SVM, and Deep learning methods are used to provide accurate risk assessments and personalized recommendations. However, creating robust models that can be applied broadly still poses a challenge. Future research should focus on integrating these models into practices. In the case of obesity, clustering algorithms like gradient boosting machines and neural networks are utilized to predict risks and pinpoint biomarkers. By leveraging dimensional data and explainable AI approaches, the accuracy of predictions can be enhanced, and the transparency of the models can be improved. Although machine learning methods are improving forecasts and customizations, for these conditions, continued work is required to tackle the issues of integration applicability and clarity in healthcare environments.

Further studies should explore in detail the elements that play a vital role in heart disease, diabetes, and obesity by conducting a thorough examination of the subject matter. To enhance the accuracy of predictions and identify biomarkers, the research should focus on combining various types of biological data, such as genomics and metabolomics, alongside proteomics. This holistic method provides insight into disease mechanisms and factors that contribute to risks. It is essential to incorporate AI techniques to improve the clarity and interpretation of ML models. Enhancing comprehension and reliability in the forecasts produced by these models will encourage healthcare professionals to integrate them into their routines. When incorporating machine learning models into healthcare systems, it is important to design interfaces that are easy for users and develop communication methods to ensure smooth implementation. Additionally, it is important to address issues like safeguarding data and preventing biases to ensure that AI is used responsibly in healthcare. Looking to these paths of exploration in the realm of machine learning in healthcare settings can lead to notable advancements that offer more accurate customized insights to handle better conditions, like heart disease, diabetes, and obesity, effectively.

References

[1] Yogesh Kumar et al., "Artificial Intelligence in Disease Diagnosis: A Systematic Literature Review, Synthesizing Framework and Future Research Agenda," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 7, pp. 8459-8486, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[2] Mahmood Safaei et al., "A Systematic Literature Review on Obesity: Understanding the Causes & Consequences of Obesity and Reviewing Various Machine Learning Approaches Used to Predict Obesity," *Computers in Biology and Medicine*, vol. 136, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[3] Mohammed Amine Makroum et al., "Machine Learning and Smart Devices for Diabetes Management: Systematic Review," *Sensors*, vol. 22, no. 5, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[4] Serena Zanelli et al., "Diabetes Detection and Management through Photoplethysmographic and Electrocardiographic Signals Analysis: A Systematic Review," *Sensors*, vol. 22, no. 13, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[5] Rahul Katarya, and Polipireddy Srinivas, "Predicting Heart Disease at Early Stages Using Machine Learning: A Survey," *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*, Coimbatore, India, pp. 302-305, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[6] Farida Mohsen et al., "A Scoping Review of Artificial Intelligence-Based Methods for Diabetes Risk Prediction," *npj Digital Medicine*, vol. 6, no. 1, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[7] Aishwarya Mujumdar, and V. Vaidehi, "Diabetes Prediction Using Machine Learning Algorithms," *Procedia Computer Science*, vol. 165, pp. 292-299, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[8] K.M. Jyoti Rani, "Diabetes Prediction Using Machine Learning," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, vol. 6, no. 4, pp. 294-305, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[9] Abid Sarwar et al., "Diagnosis of Diabetes Type-II Using Hybrid Machine Learning Based Ensemble Model," *International Journal of Information Technology*, vol. 12, no. 2, pp. 419-428, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[10] Orlando Iparraguirre-Villanueva et al., "Application of Machine Learning Models for Early Detection and Accurate Classification of Type 2 Diabetes," *Diagnostics*, vol. 13, no. 14, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[11] Harleen Kaur, and Vinita Kumari, "Predictive Modelling and Analytics for Diabetes Using a Machine Learning Approach," *Applied Computing and Informatics*, vol. 18, no. 1/2, pp. 90-100, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[12] Salleh Sonko et al., "Predicting Long-Term Type 2 Diabetes with Artificial Intelligence (AI): A Scoping Review," *Studies in Health Technology and Informatics*, vol. 305, pp. 652-655, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[13] Jayakumar Kaliappan et al., "Analyzing Classification and Feature Selection Strategies for Diabetes Prediction across Diverse Diabetes Datasets," *Frontiers in Artificial Intelligence*, vol. 7, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[14] Henock M. Deberneh, and Intaek Kim, "Prediction of Type 2 Diabetes Based on Machine Learning Algorithm," *International Journal of Environmental Research and Public Health*, vol. 18, no. 6, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[15] K. Gaurav et al., "Human Disease Prediction Using Machine Learning Techniques and Real-life Parameters," *International Journal of Engineering*, vol. 36, no. 6, pp. 1092-1098, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[16] Anita Bhatia et al., "Modeling Obesity in Complex Food Systems: Systematic Review," *Frontiers in Endocrinology*, vol. 13, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[17] Xiaobei Zhou, Lei Chen, and Hui-Xin Liu, "Applications of Machine Learning Models to Predict and Prevent Obesity: A Mini-Review," *Frontiers in Nutrition*, vol. 9, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[18] Aamir Javaid et al., "Medicine 2032: The Future of Cardiovascular Disease Prevention with Machine Learning and Digital Health Technology," *American Journal of Preventive Cardiology*, vol. 12, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[19] Dahlak Daniel Solomon et al., "Hybrid Majority Voting: Prediction and Classification Model for Obesity," *Diagnostics*, vol. 13, no. 15, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[20] Amin Gasmi, "Machine Learning and Bioinformatics for Diagnosis Analysis of Obesity Spectrum Disorders," *arXiv*, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[21] Zhaoji Fu et al., "Artificial-Intelligence-Enhanced Mobile System for Cardiovascular Health Management," *Sensors*, vol. 21, no. 3, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[22] Jessica Torres-Soto, and Euan A. Ashley, "Multi-Task Deep Learning for Cardiac Rhythm Detection in Wearable Devices," *npj Digital Medicine*, vol. 3, no. 1, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[23] Aditya Khamparia, and Karan Mehtab Singh, "A Systematic Review on Deep Learning Architectures and Applications," *Expert Systems*, vol. 36, no. 3, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[24] Francesco Goretti et al., "Deep Learning for Predicting Congestive Heart Failure," *Electronics*, vol. 11, no. 23, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[25] Balbir Singh, and Hissam Tawfik, "A Machine Learning Approach for Predicting Weight Gain Risks in Young Adults," *2019 10th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, Leeds, UK, pp. 231-234, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[26] T.M. Dugan et al., "Machine Learning Techniques for Prediction of Early Childhood Obesity," *Applied Clinical Informatics*, vol. 6, no. 3, pp. 506-520, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[27] P. Fergus et al., "A Machine Learning Approach to Measure and Monitor Physical Activity in Children to Help Fight Overweight and Obesity," *Intelligent Computing Theories and Methodologies*, pp. 676-688, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[28] Todd Lingren et al., "Developing an Algorithm to Detect Early Childhood Obesity in Two Tertiary Pediatric Medical Centers," *Applied Clinical Informatics*, vol. 7, no. 3, pp. 693-706, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[29] Muhammed Kürşad Uçar et al., "Estimation of Body Fat Percentage Using Hybrid Machine Learning Algorithms," *Measurement*, vol. 167, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[30] Tania Fernández-Navarro et al., "Exploring the Interactions between Serum Free Fatty Acids and Fecal Microbiota in Obesity through a Machine Learning Algorithm," *Food Research International*, vol. 121, pp. 533-541, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[31] G. Swapna, R. Vinayakumar, and K.P. Soman, "Diabetes Detection Using Deep Learning Algorithms," *ICT Express*, vol. 4, no. 4, pp. 243-246, 2018. [CrossRef] [Google Scholar] [Publisher Link]

Appendix

Table 1. An overview of reviewed studies

Author (s)	ML Algorithm (s)	Year	Risk Factor (s)	Findings	Drawbacks
[26]	Deep Learning	2022	Prediction of Heart Disease	The deep learning-based system effectively predicted CHF severity and progression using a medium-sized dataset of 1037 records, achieving results comparable to previous studies. The system offers valuable support for expert cardiologists by predicting disease trajectory and aiding in the design of personalized therapies and resource allocation.	The study's medium-sized dataset may limit generalizability, and the model's effectiveness could be enhanced with larger datasets and further fine-tuning. Future research is needed to explore advanced methodologies, including regression-type outputs and integration of additional data, to improve prediction accuracy and provide more detailed risk information.
[27]	Multi-Layer Perceptron Feed Forward Artificial Neural Networks (MLFFNN)	2019	Body Mass Index (BMI)	The study uses the Millennium Cohort Study data to predict teenage BMI based on earlier values, achieving over 90% accuracy with regression and neural network models. The findings highlight the effectiveness of these methods in the early detection of obesity, which is crucial for preventing related serious health conditions.	Despite achieving an impressive 93.4% prediction accuracy with the MLPFFANN algorithm, the study faces drawbacks such as the potential excessive fitting of the training data and limitations in generalizing results to different populations. Additionally, while complex algorithms were tested, they did not offer improved results and increased training times. This suggests that further refinement and exploration of alternative methods are needed for enhanced robustness and accuracy.
[28]	Random Tree (RF), J48, ID3, Naïve Bayes NB), and Bayes Trained	2015	Prediction of Obesity	The study found that the ID3 ML model, trained on CHICA data, achieved an 85% accuracy and 89% sensitivity in predicting obesity in children after age two. This model performs well and identifies strong predictors of future obesity, validating its practical use for early intervention in a clinical setting.	Although the ID3 model demonstrated high accuracy (85%) and sensitivity (89%) in predicting childhood obesity using pre-second birthday data, the study faces drawbacks such as potential limitations in model generalizability due to data quality issues and the specific nature of the CHICA dataset. Additionally, the presence of missing and erroneous values in real clinical data may affect the robustness of the model, and further research is needed to integrate and validate this model in broader clinical settings.
[29]	Artificial Neural Networks (ANN)	2015	BMI	This research study describes the use of wearable accelerometer sensors coupled with Machine Learning (ML) using an artificial neural network to provide accurate activity classification in children—up to 98.8% accuracy, 99% sensitivity and specificity – including sitting, quiet play activities typically of minimal interest to software used as PA outcomes. Another strength of this approach is that it surmounts the pitfalls associated with traditional means of measurement and establishes a robust method for assessing physical functions in natural settings.	Finally, the study becomes even more valuable when other types of machine learning are tested, and a deeper evaluation through external validation with non-interpolated databases is recommended. Second, as the analysis focused on youth only, the research findings may not be fully applicable to other age groups, and more standardization of the classifying activities is needed for broader research uniformity.

[30]	Rule-Based and ML-Based Algorithms	2016	Prediction of Obesity	According to the study, a rule-based method with accuracies of 0.895 and 0.770 at various institutions was more accurate than Machine Learning (ML) in identifying severe early childhood obesity. While ML provided flexibility in merging different aspects from the data, both methods were successful in excluding individuals who were obese because of co-occurring disorders or medication, resulting in a high-precision cohort for future clinical trials and genetic research.	Although both approaches have drawbacks, the study's rule-based algorithm fared better in identifying severe early childhood obesity than the machine-learning method. In contrast to Machine Learning (ML), which integrates features more broadly, the rule-based system may be less flexible and adaptable. On the other hand, the machine-learning technique may have difficulties in generalizing to various patient groups. Additionally, differences in EHR systems between different institutions and data quality problems may impact the algorithms' performance.
[31]	MLFFNN, SVM, and Decision Tree (DT)	2021	Body Fat Percentage (BFP)	Based on a single anthropometric measurement, the study shows that BFP may be precisely calculated utilizing hybrid machine learning techniques with few parameters. This method achieves great accuracy and practical utility in treating obesity while being less expensive than costly body analysis instruments.	With a correlation value of roughly 0.79, the suggested hybrid machine learning approach for calculating BFP using a single anthropometric measurement exhibits potential. However, because the model depends on a single measurement rather than a thorough assessment, it may have limitations regarding the generalizability of results and a potential lack of accuracy in varied populations beyond the study sample. Furthermore, the model's performance may change depending on the population's characteristics or the measuring methods used.
[32]	Decision Trees (DT)	2019	Fatty Acids	Low Eicosapentaenoic Acid (EPA) and high levels of palmitic acid, linoleic acid, gamma-linolenic acid and are characteristics of an obese-linked FFA profile and are important predictors according to this study. They discovered that in addition to Bifidobacterium, Faecalibacterium, and Bacteroides, the most important predictors of obesity were gender and serum EPA. The regression tree model linked A non-obese profile to greater levels of EPA and Bacteroides. Various obesity predictions were influenced by various microbiota.	The results highlight the potential of serum FFAs and gut microbiota as biomarkers for obesity, emphasizing the role of EPA and specific microbiota in distinguishing between obese and non-obese individuals. Future research should focus on understanding these mechanisms to develop effective strategies for obesity prevention and early detection.
[33]	SVM, CNN, Long Short Term Memory	2018	Prediction of Diabetes	The study demonstrates that deep learning architectures, specifically CNN and LSTM combined with SVM, can achieve a high accuracy of 95.7% for classifying diabetes using HRV signals from ECG data. These advanced methods offer a non-invasive and effective tool for diabetes prediction, with improvements in performance compared to previous methods.	The high accuracy achieved may not generalize across all populations or settings, especially if the input dataset size remains limited. The methodology relies on large datasets for further improvements, and the system's effectiveness in diverse or less-controlled environments may need additional validation.