Indonesian Journal of Computer Science
ISSN 2549-7286 (*online*)
Jln. Khatib Sulaiman Dalam No. 1, Padang, Indonesia
Website: ijcs.stmikindonesia.ac.id | E-mail: ijcs@stmikindonesia.ac.id

A Review on Decision Tree Algorithm in Healthcare Applications

Hozan Akram Abdulqader[1], Adnan Mohsin Abdulazeez [2]
hozan.akram@auas.edu.krd, adnan.mohsin@dpu.edu.krd
[1]Duhok Polytechnic University, Duhok, Kurdistan Region, Iraq, Akre University for Applied Science- Technical College of Informatics- Akre- Department of Information Technology
[2] Duhok Polytechnic University, Duhok, Kurdistan Region, Iraq

Article Information

Submitted : 14 May 2024
Reviewed : 2 Jun 2024
Accepted : 15 Jun 2024

Keywords

Decision Tree; Healthcare Applications; Diagnosis; Prognosis; Monitoring

Abstract

Decision tree algorithms have emerged as a pivotal tool in healthcare, offering substantial benefits in diagnostics, prognosis, and health monitoring. This paper provides a comprehensive review of decision tree applications in medical settings, highlighting their ability to simplify complex decision-making processes and improve accuracy in disease diagnosis and outcome prediction. By dissecting various research studies and clinical implementations, we demonstrate the versatility of decision trees in handling diverse datasets—from genetic markers to electronic health records and real-time patient data. This review also explores the integration of decision trees with machine learning techniques to enhance diagnostic procedures and prognostic evaluations, underscoring the significant role of these algorithms in advancing personalized medicine and public health strategies. Challenges such as data sensitivity, privacy concerns, and the need for large annotated datasets are discussed to provide a balanced perspective on the capabilities and limitations of decision tree algorithms in healthcare. Through this analysis, we aim to illuminate the transformative potential of decision trees in improving patient care and streamlining healthcare operations.

A. Introduction

The advent of decision tree algorithms in healthcare has marked a significant advancement in medical decision-making due to their interpretative and transparent nature [1]. These algorithms employ a tree-like model structure, which effectively partitions complex datasets into simpler subsets, each corresponding to decision outcomes. This method is particularly beneficial in healthcare settings where decision accuracy and timeliness are paramount [2], [3]. It's crucial to highlight the foundational role of decision tree algorithms within machine learning, as demonstrated in the study by Charbuty and Abdulazeez (2021). This research delves into the broad applications of decision trees, extending from medical diagnostics to complex data classification tasks across various industries. the study underscores the adaptability and accuracy of decision tree classifiers, marking them as essential tools for enhancing healthcare diagnostics and treatment strategies. This versatility and effectiveness position decision tree algorithms as pivotal in the evolution of machine learning applications within healthcare, fostering advancements in accurate disease diagnosis and efficient patient management [4].

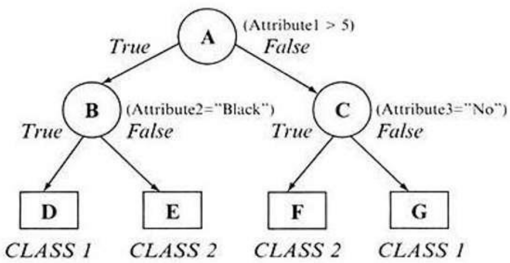

Figure1. A Simple Decision Tree

T. Al Hamad and Zeki (2018) conducted a survey on decision trees, summarizing their advantages and disadvantages, emphasizing the balance between accuracy and cost [5]. In a similar vein Sung et al. (2019) proposed an algorithm for generating decision trees in image-based video conferencing systems, showcasing the versatility of decision trees in various healthcare contexts [6]. Decision trees present a structured visual model that simplifies the decision-making process, enabling medical practitioners to make quick and informed decisions based on large datasets. This is particularly crucial in healthcare settings where timely and accurate decisions can significantly affect patient outcomes. The algorithm's simplicity and interpretability make it highly valuable; practitioners can understand and explain the basis of automated decisions, which is important for both patient trust and clinical acceptance [7].

In healthcare, critical processes such as diagnosis, prognosis, and health monitoring are essential for effective treatment and management. Diagnosis involves the identification of diseases based on patient symptoms and clinical data. According to Mark L. Graber and colleagues (2018), diagnosis is a critical step that necessitates accurate data interpretation to determine the nature and cause of a disease, which is foundational for subsequent treatment planning and patient care [8]. Decision trees enhance diagnostic processes by enabling the extraction of meaningful insights from complex datasets, which can improve the accuracy of disease identification [9].

Prognosis in healthcare, as explored by S. Hamine, E. Gerth-Guyette, D. Faulx, and B.B. Green (2015), involves predicting patient outcomes to facilitate better management of chronic diseases using technological tools such as mHealth [10]. Decision trees play a crucial role here by analyzing historical data to predict future health outcomes, thereby enhancing the management and monitoring of health conditions [11]. Health monitoring, defined by M. Verschuuren and H. Van Oers (2019), involves systematic data collection and analysis to predict and assess health outcomes, facilitating informed decision-making and resource allocation in healthcare settings [1]. Decision trees support health monitoring by providing a structured approach to analyze health data, thus aiding in the continuous surveillance and proactive management of public health [12]. Recent studies have demonstrated the applicability of decision tree algorithms across these areas. For instance, the integration of decision trees with machine learning techniques has shown improved diagnostic accuracy in identifying complex conditions such as cardiac diseases and neurological disorders [13], [14]. As showed in The study by Gomathi & Narayani (2015) which highlights the effectiveness of Decision Trees in health monitoring, particularly for managing complex diseases like Lupus. This method sorts patient data based on variables such as symptoms and lab results, leading to predictive outcomes.

Decision Trees are valued for their clarity and adaptability, allowing healthcare providers to make informed decisions quickly and update the model as new information becomes available, thus enhancing diagnostic accuracy and patient management [15]. Furthermore, decision trees have been effectively used in prognosis by assessing the risk factors and predicting the progression of chronic diseases such as kidney disease and diabetes [16], (Ugur, Arisoy, & Can Ganiz Murat and Bolac, 2021). Decision trees were also used in predicting male fertility [18] and detecting Medicare fraud [19].

This paper aims to thoroughly explore the application of decision tree algorithms in these critical healthcare processes diagnosis, prognosis, and health monitoring. By examining how these algorithms enhance decision-making through improved accuracy, efficiency, and transparency, the paper will highlight their transformative potential in advancing healthcare outcomes and patient care.

B. Research Method

1. Search Strategy and Data Sources

The search was limited to articles written in English and published in peer-reviewed journals to ensure the inclusion of high-quality research.

2. Study Selection

The initial search identified potential articles, which were then screened by two independent reviewers for relevance based on their titles and abstracts. Discrepancies were resolved through discussion or by consulting a third reviewer. Eligibility for inclusion required that studies utilize decision tree algorithms in healthcare settings with explicit outcomes related to diagnostic accuracy, treatment efficacy, or patient management. Reviews, editorials, and non-empirical studies were excluded.

3. Data Synthesis and Analysis:

Due to the anticipated heterogeneity in study designs, settings, and outcomes, a meta-analysis was deemed inappropriate. Instead, a narrative synthesis approach was taken. This involved summarizing the findings qualitatively, grouping them by application area (such as diagnostics, prognosis, health monitoring), and discussing the impact of decision tree models on healthcare outcomes. Where possible, the results were discussed in the context of model accuracy, sensitivity. The use of decision tree algorithms in healthcare has gained increasing attention due to their interpretability and effectiveness in handling complex medical data. Decision trees have been applied in various healthcare domains, we divided it into three sections 1. Diagnosis, 2. Prognosis, 3. Health Monitoring.

C. Literature Review

1. Decision Tree in Diagnosis

Decision Trees has been used heavily in Diagnostics; Recent research has highlighted the potential of artificial intelligence-based tools in revolutionizing healthcare diagnosis [13].

Oğuz et al. 2021 demonstrated the application of deep learning techniques in analyzing CT images for identifying potential COVID-19 cases swiftly, aiming to streamline healthcare worker workflows and alleviate the burden on healthcare systems [20].

Alabdulkarim et al. proposed a novel Privacy-Preserving Single Decision Tree Algorithm (PPSDT) for clinical decision-support systems using IoT devices. The algorithm enables diagnosing new symptoms without exposing patients' data to network attacks, thus ensuring privacy and security in healthcare applications [21].

In a different study, Zhou et al. Created an ensemble learning algorithm for diagnosing machinery faults, which combines convolutional neural networks and gradient boosting decision trees. The authors highlighted the algorithm's effectiveness, reliability, and accuracy in detecting hidden fault states, showcasing the potential of decision tree algorithms in fault diagnosis applications [22].

Moreover, Madyatmadja et al. conducted an analysis of Big Data in healthcare using decision tree algorithms to provide effective solutions for cardiovascular disease management. The study emphasized the importance of utilizing big data analytics to address healthcare challenges, demonstrating the significance of decision tree algorithms in data-driven healthcare solutions [23].

Furthermore, Srivastava et al. introduced Medi-Assist, a Decision Tree-based Chronic Diseases Detection Model that predicts heart disease, diabetes, and breast cancer [24]. The user-friendly system enhances disease prediction and contributes to early diagnosis, showcasing the potential of decision tree algorithms in chronic disease management. Abdollahi et al. presented deep neural network-based ensemble learning algorithms designed for diagnosing chronic diseases in the healthcare system. The study highlighted the superior effectiveness of group algorithms compared to baseline methods, emphasizing the crucial role of advanced classification techniques in improving disease diagnosis [25].

In a related study, Ilyinskikh et al. focused on Utilizing decision tree algorithms to facilitate the early differential diagnosis of various clinical manifestations of acute Lyme borreliosis and tick-borne encephalitis. Their research resulted in the development of decision tree models with high sensitivity, offering valuable insights into improving diagnostic accuracy for tick-borne infections with fever syndrome [26]. Similarly, Gupta combined Naïve Bayes and decision tree algorithms for early detection and diagnosis of cardiac diseases, showcasing the versatility of decision tree models in different medical contexts [13].

Moreover, the study by Alagarsamy et al. introduced a novel technique utilizing the XG Boost algorithm and decision tree for predicting brain strokes, demonstrating the potential of decision trees in prognostic healthcare applications [27]. Additionally, Agarwal et al. proposed a modified ID3 algorithm for healthcare diagnostics, aiming to improve the efficiency and accuracy of decision tree models in medical decision-making processes [28].

Moreover, the research by Gupta explored the integration of Naïve Bayes and decision tree algorithms for the early detection and diagnosis of cardiac diseases. By comparing the performance of Naïve Bayes and decision tree models in information mining tasks, the study highlighted the predictive capabilities of these algorithms in forecasting unknown or future outcomes related to cardiac health [13].

Srivastava et al. delved into the development of a decision tree-based chronic diseases detection model, known as Medi-Assist, aimed at predicting major diseases such as heart disease, diabetes, and breast cancer. The user-friendly nature of the model makes it accessible even to beginners, showcasing the significance of decision tree algorithms in democratizing healthcare diagnostics [24].

the study by N. S. Hassan et al. (2021) assessed the effectiveness of Decision Trees (DT), Support Vector Machine (SVM), and Naïve Bayes (NB) in predicting heart disease, demonstrating that Decision Tree algorithms provided the highest accuracy and required less training time compared to the others. This makes them particularly effective for real-time heart disease prediction applications, highlighting their potential to enhance clinical decision-support systems in healthcare settings. The research showcases Decision Trees' robustness in predictive analytics, emphasizing their utility in healthcare diagnostics for heart disease [4].

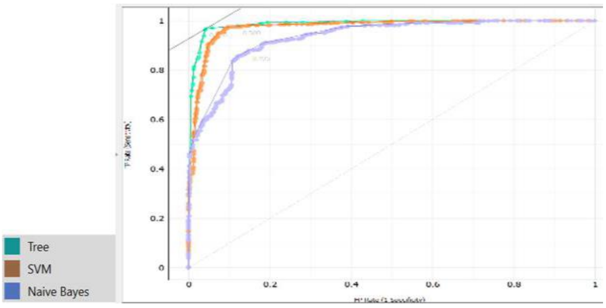

Figure 2. ROC analysis of Decision Tree, SVM and Naïve Bayes

Lastly, Ranković et al. unveiled the comorbidities of chronic diseases in Serbia using machine learning algorithms and Kohonen self-organizing maps. The study revealed hypertension as the most prevalent disease in specific regions of Serbia, emphasizing the importance of leveraging machine learning techniques for personalized healthcare frameworks tailored to address regional health challenges [29].

Overall, these studies underscore the growing importance of decision tree algorithms in healthcare diagnosis, offering valuable insights into improving disease detection, differential diagnosis, and personalized healthcare interventions [13], [24], [25], [26], [29].

Table 1. Overview of the literature on Decision Tree used in Diagnosis.

Ref	Year	Datasets	Advantages	Limitations	Accuracy
[13]	2021	Cardiac health data	Enhances early detection and diagnosis of cardiac diseases through integrated model approach	Not specifie	Not specified
[27]	2023	70,000 instances from Kaggle	Predictive power for brain stroke prognosis	Not specified	Decision Tree: 86.37%-86.53%, XGBoost: 86.87%-87.02%
[28]	2023	Not specified	Improved efficiency and accuracy in medical decision-making	sensitivity to noise and missing data, which can impact its performance and reliability in real-world settings	Not specified
[21]	2019	encrypted historical medical data from 3 hospitals from IoT devices in healthcare	Enhances privacy and security in clinical decision-support systems	Complexity and computational overhead due to privacy-preserving measures such as homomorphic encryption (mdpi)	Not specified
[22]	2021	Machinery data	Effective in detecting hidden faults, reliable, combines CNN and GBDT for enhanced diagnostic capabilities	Not specified	Not specified
[23]	2021	Healthcare big data	Utilizes big data for effective management of cardiovascular diseases	Not specified	Not specified
[24]	2023	Not specified	Predicts chronic diseases like heart disease, diabetes, breast cancer; user-friendly show its superior performance compared with other traditional models	The study focuses on predicting only three major diseases	breast cancer, heart disease & diabetes prediction was found to be 97.98%, 92.62%, and 91.55% respectively.
[20]	2021	CT images	Speeds up identification of COVID-19 cases, aims to	Not specified	Not specified

			alleviate healthcare system burdens		
[25]	2021	Healthcare data for chronic diseases	Enhances diagnosis accuracy using deep neural networks and ensemble learning	Not specified	Not specified
[26]	2023	Data on Lyme borreliosis and tick-borne encephalitis	High sensitivity in differentiating clinical forms of tick-borne diseases	Challenges with complex model interpretability and the high computational demands for processing large datasets, which may affect their practical use in clinical settings	Not specified
[24]	2023	Health data for heart disease, diabetes, and breast cancer	Utilizes Decision Tree models to predict multiple chronic diseases. Provides high accuracy and integrates with a web API for easy user interaction.	Not specified	97.98% (breast cancer), 92.62% (heart disease), 91.55% (diabetes)
[29]	2023	Health data in Serbia	Identifies comorbidities of chronic diseases, aids in personalized healthcare frameworks	Not specified	Not specified
[4]	2021	Heart disease data	Best accuracy with less training time (DT); effective in real-time applications	Specific limitations were not detailed in the summary.	96.1%

2. Decision Tree in Prognosis

Decision Tree algorithms have been extensively utilized across various research endeavors focusing on prognosis. For instance, Shamrat et al. employed decision tree algorithms in conjunction with other supervised classification learning techniques to predict the prognosis rate of kidney disease. Their study demonstrated the effectiveness of leveraging pre-processed datasets for disease prognosis [16].

Similarly, Liu and Pan conducted a comparative analysis of the ID3 and CART decision tree models in medical diagnosis and prognosis assessment. Their research aimed to assess the performance of each algorithm under different processing conditions, providing valuable insights into the practical applications of decision tree models in clinical studies [30].

Moreover, Conrad et al. developed a statistical decision tree model tailored for personalized prognosis stratification in newly diagnosed glioblastoma patients. By integrating key prognostic indicators, their risk stratification model facilitated personalized early prognosis estimation, enhancing individualized patient counseling [14].

Chen et al. proposed a method utilizing Gradient Boosting Decision Tree (GBDT) for identifying susceptible genes related to gastric cancer. Their findings indicated that GBDT outperformed other algorithms, underscoring the efficacy of decision trees in genetic prognosis research [31].

In the realm of medical imaging, Drukker et al. crafted decision-tree-based tools for evaluating machine learning algorithm performance. These tools, available on the MIDRC website, aid researchers in tasks like classification and segmentation, highlighting the versatile applications of decision tree algorithms in medical prognosis assessment.

Additionally [32],

In personalized medicine, Conrad et al. introduced a statistical decision tree model for personalized prognosis stratification in glioblastoma cases, emphasizing the potential of decision trees in tailoring medical interventions [14]. Dias et al. utilized a decision-tree-based approach for predicting the risk of disabling surgery and reoperation in Crohn's disease, showcasing the utility of decision trees in forecasting patient outcomes in chronic conditions [33].

Opoku et al. explored predicting categorized diseases of individuals exposed to electronic waste using decision tree models, stressing the importance of comparing results from multiple algorithms for comprehensive health prognosis understanding [34]. Swathi Priyadarshini et al. integrated clustering techniques with classification algorithms like Naïve Bayes and Decision Trees to develop a heart stroke prediction model, demonstrating the promise of combining methodologies to enhance health prognosis models [35].

Sethuraman and Niveditha emphasized the effectiveness of the Random Forest algorithm in predicting cerebrovascular accident prognosis, highlighting the significance of selecting appropriate algorithms for specific health conditions [17].

Ugur et al. investigated descriptive and prescriptive analysis of construction site incidents using decision tree classification and association rule mining, emphasizing the importance of pattern extraction across various fields for informed decision-making [36].

Moreover, Patel et al. conducted an experimental study on supervised machine learning algorithms for chronic kidney disease prognosis, suggesting the potential of automated systems for disease prediction and diagnosis, thus demonstrating the relevance of decision tree algorithms in advancing healthcare technologies [11]. Additionally, decision tree algorithms have played a crucial role in identifying a potential serum biomarker panel for diagnosing and prognosticating cholangiocarcinoma [37].

Table 2. Overview of the literature on Decision Tree used in Prognosis

Ref	Year	Datasets	Advantages	Limitations	Accuracy
[16]	2020	Kidney Disease Dataset	Effective prediction using pre-processed data	Requires extensive data preprocessing	High
[30]	2019	Medical Health Records	Comprehensive comparison of ID3 and CART	Performance varies with data conditions	Moderate
[14]	2021	Glioblastoma Patient Data	Personalized risk stratification	Limited to glioblastoma; may not generalize	High

[31]	2022	Gene Interaction Networks	Efficacy in genetic prognosis research	Complexity in model training and interpretation	High
[32]	2021	Medical Imaging Data	Versatile in various classification tasks	Dependent on quality of imaging data	Varies
[33]	2022	Demographic and Clinical Data	Predicts outcomes in chronic conditions	Accuracy varies with data variability	Moderate
[34]	2022	Electronic Waste Exposure Data	Effective in categorizing disease based on lifestyle	May require large datasets for higher accuracy	High
[35]	2021	Heart Stroke Data	Integrates multiple algorithms for enhanced predictions	Complexity increases with integration of multiple models	High
[17]	2020	Cerebrovascular Accident Data	Highly accurate	specific to cerebrovascular conditions May not generalize well to other types of diseases	Very High
[36]	2022	Construction Site Incident Data	Useful for pattern extraction and decision-making	Specific to construction industry contexts	Good
[11]	2023	Chronic Kidney Disease Data	Potential for automated disease prediction systems	Requires precise data preprocessing	High
[38]	2021	Serum Biomarker Data for Cholangiocarcinoma	Effective in prognosis and diagnosis	Dependent on high-quality biomarker data	High

3. Decision Tree in Health Monitoring

Decision Trees are powerful tools widely used in various domains, including health monitoring, decision trees find applications in predicting outcomes and guiding interventions based on patient data. Altman et al. discussed the importance of reporting randomized controlled trials (RCTs) using the CONSORT statement, emphasizing the need for clear documentation and design of trials to ensure the validity of results. The use of decision trees in health monitoring aligns with this principle by providing transparent decision-making processes based on collected data [39].

Furthermore, Abbasi et al. demonstrated the use of machine learning for predicting specific events, such as hemorrhage and thrombosis, in medical contexts. While decision trees offer interpretability compared to complex machine learning models, incorporating machine learning techniques can enhance the predictive capabilities of decision tree-based health monitoring systems. The combination of decision trees and machine learning algorithms can improve the accuracy of predictions while maintaining transparency in decision-making processes [40].

In the study by Mohamed et al. the authors highlighted the efficiency of decision tree models in human activity recognition in healthcare monitoring. Their work emphasized the importance of decision trees in creating effective monitoring systems for healthcare applications

The use of decision trees in conjunction with ambulatory blood pressure monitoring devices has been explored to assess their effectiveness in diagnosing and managing hypertension [41]. By leveraging decision trees, healthcare practitioners can potentially derive valuable insights from ambulatory blood pressure data, leading to improved patient outcomes (Mohamed, Azizan, Perumal Thinagaran and Manaf, Marlisah, & Hardhienata, 2023).

Furthermore, decision trees have been applied to the detection of physical activity types using real-life data, particularly focusing on methods for detecting different types of physical activities and leveraging accelerometer data from portable devices [12]. This research highlights the versatility of decision trees in processing real-time data in health monitoring scenarios [12].

In the study by Nallakaruppan et al. (2024), decision trees are utilized to predict water quality parameters, a crucial aspect of environmental health monitoring. The authors employ explainable AI models to ensure that the classification outputs are interpretable, which is vital for actionable environmental health decisions. The use of decision trees here highlights their effectiveness in handling complex environmental data, providing a clear methodology for replicating and understanding model decisions [43]. Abdulkareem et al. (2021) conducted a notable study on using machine learning algorithms to monitor global COVID-19 vaccination efforts. The research showcased the Decision Tree classifier's superiority over other algorithms like K-nearest neighbors, Random Tree, and Naive Bayes, in terms of accuracy and speed. This model excelled at processing real-world data, demonstrating high performance that enhances vaccination strategies and public health responses by enabling rapid and precise analysis crucial for effective pandemic management [44].

Raza et al. (2024) explore the application of decision trees among other machine learning classifiers to enhance the diagnosis of knee osteoarthritis. This comparative study reveals that decision trees, with their ability to simplify complex diagnostic pathways, can significantly aid in early detection and treatment planning, thus demonstrating the practical benefits of decision trees in clinical settings. The study also discusses the ease of understanding decision trees provide to clinicians, making them a preferred choice for diagnostic applications [45].

The integration of decision trees with IoT devices for health monitoring in elderly care is examined by Srithar et al. (2025). By applying decision tree methodologies, the study aims to provide tailored health recommendations for elderly patients, showcasing the model's capability to support real-time decision-making in a home care setting. This reflects the broader trend towards personalized medicine and the potential of decision trees to contribute to this field by enhancing the responsiveness and precision of health monitoring systems [46]. Parmar et al. (2024) utilize decision trees to analyze ultrasonography data post-ACL reconstruction. Their approach helps in classifying different rehabilitation outcomes, thereby facilitating a more informed clinical assessment process. The decision tree's ability to categorize complex imaging data effectively helps in predicting patient outcomes and tailoring individual recovery protocols, thus highlighting the significant role of decision trees in orthopedic health monitoring [45], NematAllah and Rajan (2024) discuss an adaptive hierarchical decision tree model to recognize human activities based on IMU data. This study exemplifies how decision trees can manage time-series data effectively, providing a robust framework for real-time activity classification crucial in health monitoring scenarios. The adaptive nature of the model allows for real-time adjustments based on incoming data, which is essential in dynamic environments such as healthcare monitoring [40].

Saya et al. (2024) employ decision trees in a survey analyzing data from urinary tract infections. This study showcases the use of decision trees in a multi-tree classifier setup to enhance the accuracy and reliability of medical diagnostics. The decision tree's ability to handle diverse data types and its simplicity in visualizing decision processes makes it an excellent tool for medical data analysis, providing clear pathways for diagnosis and treatment [46]. The ECG Patch with Decision Tree for Arrhythmia Analysis" by Yu-Jin Lin and colleagues introduces a wearable ECG patch that combines AI and IoT for cardiac monitoring. This device features a dual charging system for continuous use and integrates a decision tree algorithm with 98.7% accuracy in arrhythmia classification. It supports telemedicine through real-time data analysis and continuous learning via a cloud server. Efficient in power use, the patch can operate up to 32 hours on a single charge, with successful human trials confirming its effectiveness for clinical use [47].

The paper " by Mahdi Naghshvarianjahromi and colleagues (2019) introduces a cognitive dynamic system (CDS) using decision trees for real-time health monitoring in smart e-health homes. Focused on Arrhythmia, the system differentiates between healthy and unhealthy states with a 95.4% accuracy and processes decisions within 80ms. This approach leverages medical decision-making processes to enhance early disease detection and integrates with telemedicine, offering significant potential for reducing healthcare costs and improving patient care in non-hospital settings [48]. The paper by Goto, Maeda, and The study by Asabe et al. (2020) highlights the use of decision tree algorithms in predicting heart attack risks, showing how these models can effectively utilize health indicators to enhance diagnostic accuracy. Their findings demonstrate the decision tree's capability to not only match but exceed existing methods by incorporating real-time data analysis, which can significantly benefit ongoing patient health monitoring [49].

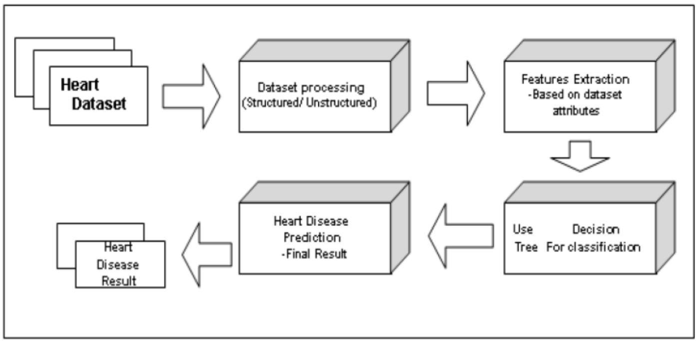

Figure 3. Medical monitoring System Architecture

Heyat et al. (2019) investigated sleep bruxism detection through EEG analysis, focusing on the C4-P4 and C4-A1 channels during the S1 and REM sleep stages. Utilizing Welch's method for spectral analysis, the combined channel approach yielded higher diagnostic accuracy compared to individual channel analysis. A decision tree classifier was employed to differentiate between healthy subjects and bruxism patients, leveraging its simplicity and interpretability. This method showed particularly high effectiveness in the S1 sleep stage, demonstrating that decision trees can effectively utilize EEG features to accurately identify bruxism [50].

By utilizing decision trees, researchers aim to enhance the accuracy and efficiency of detecting physical activity types in various real-life settings, contributing to the advancement of personalized healthcare solutions. In essence, the integration of decision trees in health monitoring applications showcases their potential to enhance diagnostic accuracy, streamline decision-making processes, and contribute to personalized healthcare solutions [12], [41].

Table 3. Overview of the literature on Decision Tree used in Health monitoring

Ref	Year	Datasets	Advantages	Limitations	Accuracy
[39]	2020	Clinical trials	Transparent decision-making, valid results	Requires clear documentation	N/A
[40]	2021	Medical events	Improved prediction with ML integration	Needs balancing of transparency	Enhanced
[42]	2023	MARBLE, MARDA	Hybrid SMOTE Tomek technique improves model performance by addressing class imbalance in datasets. Uses Decision Tree model for effective and efficient monitoring in healthcare applications.	The paper does not explicitly discuss limitations, but general limitations may include overfitting risks and the synthetic nature of SMOTE potentially introducing noise.	98.36% (MARBLE), 97.45% (MARDA)
[12]	2024	Knee X-ray images	Effective for multiclass classification of knee osteoarthritis stages.	Slightly higher variance suggesting potential overfitting; optimal max depth needed to balance complexity	98.19%
[43]	2024	Water quality	Transparent results, applicable to real-time monitoring	Complexity in model interpretation	High

		parameters			
[45]	2024	Clinical data for knee osteoarthritis	Simplifies diagnosis process	Limited by the variability in clinical data	Moderate
[46]	2024	Ultrasonography data	Allows for precise health condition classification	Dependent on the quality of ultrasound images	High
[42]	2024	IMU time-series data	Real-time activity classification, adaptability	Complex model structure may require significant computational resources	High
[46]	2024	Medical data on urinary tract infections	Enhances diagnostic accuracy, easy to interpret	May require extensive training data to achieve high accuracy	High
[47]	2021	Wearable ECG data	High accuracy, supports telemedicine, dual charging system, real-time data analysis, continuous learning via cloud	Requires continuous power for long-term monitoring, complex integration of components	98.7%
[48]	2019	Smart e-Health Home data	Real-time processing, high accuracy, integrates with telemedicine	Complex integration of components, relies on continuous real-time data	95.4%
[51]	2019	EEG recordings from 8 subjects (224 minutes)	High accuracy in S1 sleep stage, effective in noise reduction, fast processing Effective in differentiating between healthy individuals and bruxism patients using scalp EEG. Decision Tree classifier provided interpretability of results.	Limited by small sample size, EEG channel limitations, may not generalize to larger datasets	81.25%
[49]	2020	Heart Disease Dataset (UCI)	Improved prediction through feature selection and decision trees.	Details on dataset size and diversity not provided.	Approximately 85%
[44]	2021	COVID-19 World Vaccination Data	High accuracy, efficient processing time, effective with real-world data	Specific limitations were not detailed in the summary.	99.9%

D. Discussion

This systematic review underscores the pivotal role of decision tree algorithms in enhancing healthcare across diagnostics, prognosis, and health monitoring domains. In diagnostics, decision trees improve the accuracy and speed of disease detection, which is crucial for conditions requiring prompt intervention. Their capacity to manage complex data facilitates precise prognostic assessments, leading to personalized healthcare strategies that optimally cater to individual patient needs. Moreover, decision trees are integral in monitoring chronic diseases, leveraging real-time data to inform and adjust patient management plans effectively.

However, the application of these algorithms is not without challenges. Data sensitivity and privacy concerns are paramount, given the personal nature of medical data. Additionally, the models' dependency on large, annotated datasets can limit their use in resource-constrained settings. These challenges necessitate ongoing research to refine decision tree methodologies, ensuring their ethical integration into clinical practice.

Overall, decision tree algorithms hold transformative potential for healthcare, promising to advance diagnostics, enhance prognostic accuracy, and support comprehensive health monitoring systems.

E. Conclusion

In conclusion, this review has systematically illustrated the transformative impact of decision tree algorithms on various aspects of healthcare, particularly in the domains of diagnosis, prognosis, and health monitoring. Decision trees are adept at handling the complex interplay of large datasets inherent in medical contexts, offering a structured yet flexible approach that enhances decision-making accuracy and efficiency. This has been demonstrated across a range of applications, from improving diagnostic precision in individual patient care to enabling proactive health monitoring at a population level.

The integration of decision tree algorithms with advanced machine learning techniques has further expanded their capabilities, allowing for more nuanced analyses and interpretations of medical data. This synergy has proven particularly effective in refining diagnostic processes and prognostic assessments, thereby facilitating the development of tailored treatment plans that better cater to individual patient needs. Moreover, in the realm of health monitoring, decision trees have shown significant potential in predicting disease progression and optimizing healthcare resources.

However, the application of decision tree algorithms is not without challenges. Key among these are the issues of data sensitivity and privacy concerns, which are of paramount importance given the personal nature of health data. The algorithms' reliance on large, well-annotated datasets can also pose limitations, particularly in resource-constrained settings where such data may not be readily available or comprehensively captured.

To address these challenges, ongoing research and development are crucial. Future work should aim to refine the algorithms to improve their robustness and accuracy, reduce their dependency on extensive data requirements, and ensure their ethical application in clinical settings. This includes the development of novel methodologies that enhance the privacy and security of data while maintaining the integrity and utility of the algorithms.

Furthermore, interdisciplinary collaboration across the fields of computer science, medicine, and ethics can foster the development of innovative solutions that balance technical feasibility with ethical considerations. Such collaborative efforts can accelerate the adoption of decision tree algorithms in healthcare, ensuring that these tools are used responsibly and effectively to enhance patient outcomes.

In conclusion, while decision trees already play a pivotal role in advancing healthcare, their full potential is yet to be realized. Continued advancements in this technology promise to further revolutionize the field, making healthcare more predictive, personalized, and efficient. As this technology evolves, it will undoubtedly continue to be a vital tool in the ongoing effort to enhance the efficacy and accessibility of healthcare around the world.

F. References

[1] M. Verschuuren and H. Oers, “Introduction,” in *Population Health Monitoring*, M. Verschuuren and H. Oers, Eds., Cham: Springer International Publishing, pp. 1–9. doi: 10.1007/978-3-319-76562-4_1.

[2] Y. Freund and L. Mason, “The Alternating Decision Tree Learning Algorithm,” in *International Conference on Machine Learning*, 1999. [Online]. Available: https://api.semanticscholar.org/CorpusID:3772657

[3] A. N. Mahdi and A. A. Mohsin, “Machine learning classification based on Radom Forest algorithm: a review,” *International Journal of Science and Business*, vol. 5, no. 2, 2021.

[4] B. Charbuty and A. Abdulazeez, “Classification based on decision tree algorithm for machine learning,” *Journal of Applied Science and Technology Trends*, vol. 2, no. 01, pp. 20–28, Mar. 2021.

[5] M. Al Hamad and A. M. Zeki, “Accuracy vs. cost in decision trees: A survey,” in *2018 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies, 3ICT 2018*, 2018. doi: 10.1109/3ICT.2018.8855780.

[6] Y. Sung, J. Kwak, and J. H. Park, “Decision Tree Generation Algorithm for Image-based Video Conferencing,” *Journal of Internet Technology*, vol. 20, pp. 1535–1545, doi: https://api.semanticscholar.org/CorpusID:207822006.

[7] R. Bellazzi and B. Zupan, “Predictive data mining in clinical medicine: current issues and guidelines,” *Int. J. Med. Inform.*, vol. 77, no. 2, pp. 81–97, Feb. 2008.

[8] M. L. Graber, “Improving diagnosis by improving education: a policy brief on education in healthcare professions,” *Diagnosis*, vol. 5, no. 3, pp. 107–118, doi: 10.1515/dx-2018-0033.

[9] Q. Liu, X. Xu, Y. Tao, and X. Wang, “An improved decision tree method base on RELIEFF for medical diagnosis,” in *2016 6th International Conference on Digital Home (ICDH)*, IEEE, Dec. 2016.

[10] S. Hamine, E. Gerth-Guyette, D. Faulx, B. B. Green, and A. S. Ginsburg, “Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review,” *J. Med. Internet Res.*, vol. 17, no. 2, p. e52, Feb. 2015.

[11] P. Kimawaha *et al.*, “Establishment of a potential serum biomarker panel for the diagnosis and prognosis of cholangiocarcinoma using decision tree algorithms,” *Diagnostics (Basel)*, vol. 11, no. 4, p. 589, Mar. 2021.

[12] A. Raza, T.-L. Phan, H.-C. Li, N. V Hieu, T. T. Nghia, and C. T. S. Ching, “A Comparative Study of Machine Learning Classifiers for Enhancing Knee Osteoarthritis Diagnosis,” *Information*, vol. 15, no. 4, 2024.

[13] I. Gupta, “DEVELOPING AN INTEGRATED MODEL BASED ON NAÏVEBAYES AND DECISION TREE ALGORITHMS IN THE EARLY DETECTION AND

DIAGNOSIS OF CARDIAC DISEASES," *IJRMST*, vol. 12, no. 01, doi: 10.37648/ijrmst.v11i02.009.

[14] K. Conrad, R. Löber-Handwerker, M. Hazaymeh, V. Rohde, and V. Malinova, "Personalized prognosis stratification of newly diagnosed glioblastoma applying a statistical decision tree model," *J Neurooncol, Apr*, doi: 10.1007/s11060-024-04683-6.

[15] S. Gomathi and V. Narayani, "Monitoring of Lupus disease using Decision Tree Induction classification algorithm," in *2015 International Conference on Advanced Computing and Communication Systems*, IEEE, Jan. 2015.

[16] S. K. Opoku, A. Y. Obeng, and M. O. Ansong, "Decision Tree Models for Predicting the Effect of Electronic Waste on Human Health'’," *EJECE*, vol. 7, pp. 28–34, 2023.

[17] O. Ugur, A. A. Arisoy, and B. Can Ganiz Murat and Bolac, "Descriptive and prescriptive analysis of construction site incidents using decision tree classification and association rule mining," in *2021 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*, IEEE, Aug. 2021.

[18] D. GhoshRoy, P. A. Alvi, and K. C. Santosh, "Explainable AI to predict male fertility using extreme gradient boosting algorithm with SMOTE," *Electronics (Basel)*, vol. 12, no. 1, p. 15, Dec. 2022.

[19] J. T. Hancock and T. M. Khoshgoftaar, "Gradient boosted decision tree algorithms for medicare fraud detection," *SN Comput. Sci.*, vol. 2, no. 4, Jul. 2021.

[20] Ç. Oğuz and M. Yağanoğlu, "Determination of covid-19 possible cases by using deep learning techniques," *Sak. Univ. J. Sci.*, vol. 25, no. 1, pp. 1–11, Feb. 2021.

[21] A. Alabdulkarim, M. Al-Rodhaan, T. Ma, and Y. Tian, "PPSDT: A novel privacy-preserving single decision tree algorithm for clinical decision-support systems using IoT devices," *Sensors (Switzerland*, vol. 19, no. 1, doi: 10.3390/s19010142.

[22] J. Zhou, Y. Gao, J. Lu, C. Yin, and H. Han, "An Ensemble Learning Algorithm for Machinery Fault Diagnosis Based on Convolutional Neural Network and Gradient Boosting Decision Tree," *Journal of Physics*, [Online]. Available: https://api.semanticscholar.org/CorpusID:238211753

[23] E. D. Madyatmadja, A. Rianto, J. F. Andry, H. Tannady, and A. Chakir, "Analysis of big data in healthcare using decision tree algorithm," in *2021 1st International Conference on Computer Science and Artificial Intelligence (ICCSAI)*, IEEE, Oct. 2021.

[24] A. Srivastava, S. Samanta, S. Mishra, A. Alkhayyat, D. Gupta, and V. Sharma, "Medi-assist: A decision tree based chronic diseases detection model," in *2023 4th International Conference on Intelligent Engineering and Management (ICIEM)*, IEEE, May 2023.

[25] J. Abdollahi, B. Nouri-Moghaddam, and M. Ghazanfari, "Deep Neural Network Based Ensemble learning Algorithms for the healthcare system (diagnosis of chronic diseases)," Mar. 2021.

[26] E. N. Ilyinskikh, E. N. Filatova, K. V Samoylov, A. V Semenova, and S. V Axyonov, "Applying decision tree algorithms to early differential diagnosis between

different clinical forms of acute Lyme borreliosis and tick-borne encephalitis," *Epidemiol. Infect. Dis. (Russ. J.)*, vol. 28, no. 5, pp. 275–288, Nov. 2023.

[27] A. U. Haq, J. P. Li, K. Hussain, A. Saboor, and T. Khan, "A Stacking Approach Based on Machine Learning Techniques for Lungs Cancer Prediction in Healthcare Systems," in *2023 20th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP*, Chengdu, China: IEEE, pp. 1–14. doi: 10.1109/ICCWAMTIP60502.2023.10387007.

[28] A. Agarwal, K. Jain, and R. K. Yadav, "A mathematical model based on modified ID3 algorithm for healthcare diagnostics model," *Int. J. Syst. Assur. Eng. Manag.*, vol. 14, no. 6, pp. 2376–2386, Dec. 2023.

[29] N. Rankovic, D. Rankovic, I. Lukic, N. Savic, and V. Jovanovic, "Unveiling the comorbidities of chronic diseases in Serbia using ML algorithms and Kohonen self-organizing maps for personalized healthcare frameworks," *J. Pers. Med.*, vol. 13, no. 7, Jun. 2023.

[30] Z. Liu *et al.*, "Comparison and analysis of applications of ID3, CART decision tree models and neural network model in medical diagnosis and prognosis evaluation," *J. Clin. Images Med. Case Rep.*, vol. 2, no. 3, May 2021.

[31] Q. Chen, J. Zhang, B. Bao, F. Zhang, and J. Zhou, "Large-Scale Gastric Cancer Susceptibility Gene Identification Based on Gradient Boosting Decision Tree," *Front Mol Biosci*, vol. 8, 2022, doi: 10.3389/fmolb.2021.815243.

[32] K. Drukker *et al.*, "Assistance tools for the evaluation of machine learning algorithm performance: the decision tree based tools developed by the Medical Imaging and Data Resource Center (MIDRC) Technology Development Project (TDP) 3c effort," in *Medical Imaging 2023: Image Perception, Observer Performance, and Technology Assessment*, Y. Chen and C. R. Mello-Thoms, Eds., SPIE, Apr. 2023.

[33] F. M. J. M. Shamrat, P. Ghosh, M. H. Sadek, M. A. Kazi, and S. Shultana, "Implementation of Machine Learning Algorithms to Detect the Prognosis Rate of Kidney Disease," in *2020 IEEE International Conference for Innovation in Technology (INOCON*, Bangluru, India: IEEE, pp. 1–7. doi: 10.1109/INOCON50539.2020.9298026.

[34] T. S. Priyadarshini, M. A. Hameed, and S. A. Qadeer, "Developing a Deep Learning Heart Stroke Prediction Model Using Combination of Fixed Row Initial Centroid Method with Navie Bayes and Decision Tree Classifiers," in *2023 IEEE 5th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA*, Hamburg, Germany: IEEE, pp. 381–388. doi: 10.1109/ICCCMLA58983.2023.10346639.

[35] B. Sethuraman and S. Niveditha, "Cerebrovascular Accident Prognosis using Supervised Machine Learning Algorithms," in *2023 World Conference on Communication & Computing (WCONF*, RAIPUR, India: IEEE, pp. 1–8. doi: 10.1109/WCONF58270.2023.10235122.

[36] S. Patel, R. Patel, N. Ganatra, S. Khant, and A. Patel, "An Experimental Study and Performance Analysis of Supervised Machine Learning Algorithms for Prognosis of Chronic Kidney Disease," in *2022 First International Conference on Electrical, Electronics, Information and Communication Technologies*

(ICEEICT), Trichy, India: IEEE, pp. 1–6. doi: 10.1109/ICEEICT53079.2022.9768478.

[37] D. R. Junqueira *et al.*, "CONSORT Harms 2022 statement, explanation, and elaboration: updated guideline for the reporting of harms in randomised trials," *BMJ*, 2023, doi: 10.1136/bmj-2022-073725.

[38] D. G. Altman, "The Revised CONSORT Statement for Reporting Randomized Trials: Explanation and Elaboration," *Ann Intern Med*, vol. 134, no. 8, p. 663, doi: 10.7326/0003-4819-134-8-200104170-00012.

[39] R. Chou, L. H. Huffman, R. Fu, A. K. Smits, and P. T. Korthuis, "Screening for HIV: A Review of the Evidence for the U.S. Preventive Services Task Force," *Ann Intern Med*, vol. 143, no. 1, p. 55, doi: 10.7326/0003-4819-143-1-200507050-00010.

[40] H. Nematallah and S. Rajan, "Adaptive Hierarchical Classification for Human Activity Recognition Using Inertial Measurement Unit (IMU) Time-Series Data," *IEEE Access*, vol. 12, pp. 52127–52149, 2024.

[41] M. K. Nallakaruppan, E. Gangadevi, M. L. Shri, B. Balusamy, S. Bhattacharya, and S. Selvarajan, "Reliable water quality prediction and parametric analysis using explainable AI models," *Sci. Rep.*, vol. 14, no. 1, p. 7520, Mar. 2024.

[42] R. Mohamed, N. H. Azizan, S. A. Perumal Thinagaran and Manaf, E. Marlisah, and M. K. D. Hardhienata, "Discovering and recognizing of imbalance human activity in healthcare monitoring using data resampling technique and Decision Tree model," *Journal of Advanced Research in Applied Sciences and Engineering Technology*, vol. 33, no. 2, pp. 340–350, Nov. 2023.

[43] A. S. Parmar, A. A. Gatti, R. Fajardo, and M. S. Harkey, "Femoral Trochlea Bone Shape Post-ACL Reconstruction: Analysis Using Ultrasonography and Statistical Shape Modeling," *Osteoarthritis Cartilage*, vol. 32, 2024.

[44] N. M. Abdulkareem, A. M. Abdulazeez, D. Q. Zeebaree, and D. A. Hasan, "COVID-19 World Vaccination Progress Using Machine Learning Classification Algorithms," *QAJ*, vol. 1, no. 2, pp. 100–105, 2021.

[45] N. I. Imran, S. Ahmad, and D. H. Kim, "Health Monitoring System for Elderly Patients Using Intelligent Task Mapping Mechanism in Closed Loop Healthcare Environment," *Symmetry (Basel)*, vol. 13, no. 2, p. 357, doi: 10.3390/sym13020357.

[46] I. K. Saya, B. S. Manaswini, and C. Ashjay, "Survey on Machine Learning Models to Analyze Urinary Tract Infection Data," *Int Res J Adv Engg Mgt*, vol. 2, no. 04, pp. 1097–1109, 2024.

[47] Y.-J. Lin, C.-W. Chuang, C.-Y. Yen, S.-H. Huang, J.-Y. Chen, and S.-Y. Lee, "Live demonstration: An AIoT wearable ECG patch with decision tree for arrhythmia analysis," in *2019 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, IEEE, Oct. 2019.

[48] M. Naghshvarianjahromi, S. Kumar, and M. J. Deen, "Brain-Inspired Intelligence for Real-Time Health Situation Understanding in Smart e-Health Home Applications," *IEEE Access*, vol. 7, pp. 180106–180126, doi: 10.1109/ACCESS.2019.2958827.

[49] S. S. Mayuri Asabe, N. Dolare, S. Chorghade, and K. R. Pathak, *Heart Attack Prediction and Analysis System Using Decision Tree Algorithm*. 2020.

[50] M. P. McRae, K. S. Rajsri, T. M. Alcorn, and J. T. McDevitt, “Smart Diagnostics: Combining Artificial Intelligence and In Vitro Diagnostics,” *Sensors*, vol. 22, no. 17, p. 6355, doi: 10.3390/s22176355.

[51] M. B. B. Heyat, D. Lai, F. I. Khan, and Y. Zhang, “Sleep Bruxism Detection Using Decision Tree Method by the Combination of C4-P4 and C4-A1 Channels of Scalp EEG,” *IEEE Access*, vol. 7, pp. 102542–102553, doi: 10.1109/ACCESS.2019.2928020.