Available online www.unicrossjournals.com

UNICROSS JOURNAL OF SCIENCE AND TECHNOLOGY, UJOST

RESEARCH ARTICLE VOL. 4(4) DECEMBER, 2025 ISSN:2814-2233

Date Approved December 31, 2025 Pages 93 – 99

USING DECISION TREE ALGORITHM AND RULE-BASED SYSTEMS FOR MEDICAL DIAGNOSIS

Dr. Prince O. Ana

Department of Computer Science

University of Cross River State, Calabar.

Abstract

Clinical Decision Support Systems (CDSS) enhance diagnostic precision, improve treatment safety, and ensure adherence to clinical guidelines by tailoring recommendations to individual patients. This paper proposes a hybrid framework that integrates rule-based reasoning capturing guidelines, contraindications, and dosage rules with decision tree learning, which models complex, nonlinear relationships in patient data. Decision trees uncover diagnostic patterns from large datasets, offering interpretable decision paths, while rule-based systems encode expert knowledge into explicit "if–then" rules, ensuring transparency and consistency. The proposed architecture: (1) encodes expert knowledge into a structured rule base; (2) learns decision trees from heterogeneous datasets to generate preliminary assessments; (3) reconciles outputs through a coordination module that manages uncertainty and produces clinician-friendly rationales; and (4) evaluates performance using simulated and real-world scenarios. A brief case study illustrates practical implementation. Findings highlight the complementary nature of decision trees and rule-based systems, underscoring their potential to improve diagnostic accuracy and patient outcomes in modern healthcare.

Keywords: Decision Tree, Rule-Based Systems, Medical Diagnosis, AI in Healthcare

1.0 Introduction

Accurate diagnosis is central to effective healthcare, directly influencing treatment quality and patient outcomes. However, the increasing complexity of modern medicine driven by vast amounts of clinical data and rapidly evolving medical knowledge poses significant challenges for clinicians. Traditional diagnostic practices, though grounded in professional expertise, remain vulnerable to human error, cognitive bias, and variability in judgment.

Artificial intelligence (AI) has emerged as a transformative tool in medical diagnostics, supporting healthcare providers by analyzing large datasets, reducing diagnostic uncertainty, and promoting evidence-based decision-making. Among AI techniques, decision tree models and rule-based expert systems are particularly influential due to their clarity, interpretability, and adaptability to complex diagnostic scenarios (Quinlan, 1986; Han et al., 2012; Shortliffe, 1976; Sharma & Rani, 2020; Kononenko, 2001).

Decision trees classify diseases by learning patterns from structured patient data such as symptoms, laboratory results, and clinical indicators. Their strength lies in producing transparent, stepwise decision paths that clinicians can readily validate. In contrast, rule-based expert systems encode medical knowledge into explicit "if–then" rules, ensuring consistent application of guidelines and expert reasoning.

Integrating these two approaches offers a hybrid framework that not only improves diagnostic accuracy but also enhances scalability in managing diverse patient populations. This paper examines the theoretical foundations, practical applications, and comparative strengths of decision tree algorithms and rule-based systems, highlighting their complementary roles in advancing modern healthcare.

2.0 Methodology

This study adopts a qualitative review methodology to critically examine the use of decision tree algorithms and rule-based systems in medical diagnosis. The research design compares and synthesizes findings from recent literature, focusing on theoretical foundations, implementation strategies, and clinical applications.

A structured literature search was conducted in PubMed, IEEE Xplore, ScienceDirect, and Google Scholar. Selection criteria included peer-reviewed journal articles, case studies, and comprehensive reviews published between 1980 and 2023. Foundational works such as Quinlan (1986), Shortliffe (1976), and Kononenko (2001) were included to establish historical context, while more recent studies (e.g., Han et al., 2012; Sharma & Rani, 2020) provided insights into current technological advancements.

The methodology comprised three stages:

1. **Comparative analysis** of decision tree and rule-based systems, evaluating their strengths, limitations, and suitability for diagnostic tasks.
2. **Illustrative case study** demonstrating integration of both techniques in a clinical setting, highlighting practical considerations and synergistic potential.
3. **Focused assessment** of data quality, interpretability, and scalability, examining their impact on the effectiveness and adoption of AI-driven diagnostic tools.

This approach provides a comprehensive understanding of how decision trees and rule-based systems contribute to diagnostic accuracy, clinical decision support, and system efficiency, offering a foundation for future research and implementation strategies.

Medical Diagnosis Using Decision Tree Algorithms

Decision trees are supervised machine learning models that classify data by recursively splitting it into branches based on feature values, with final outcomes represented at leaf nodes. In medical diagnosis, they are widely applied for disease classification, risk prediction, and clinical decision support. For example, decision trees can distinguish between benign and malignant tumors using patient attributes such as age, laboratory results, and symptoms.

A key strength of decision trees lies in their interpretability: clinicians can trace the decision path to understand the rationale behind each diagnosis. They are capable of handling both categorical and numerical data, making them suitable for heterogeneous medical datasets. Advanced ensemble methods, such as Random Forests and Gradient Boosted Trees, further enhance predictive accuracy by combining multiple decision trees.

Nonetheless, decision trees are prone to overfitting, particularly when trained on small or noisy datasets, and their performance depends heavily on the quality and representativeness of training data. Despite these limitations, they remain widely adopted in healthcare due to their transparency, adaptability, and ease of use. See Figure 1 Hybrid Decision Tree + Rule-Base flowchart

How this hybrid works:

- Use the **Decision Tree** for high-level, mutually exclusive splits
- When a branch requires frequently updated logic, hand off to the **Rule Base** (IF–THEN rules like compliance checks, thresholds, exceptions).
- This keeps the tree compact while the rule base remains flexible and easy to maintain.

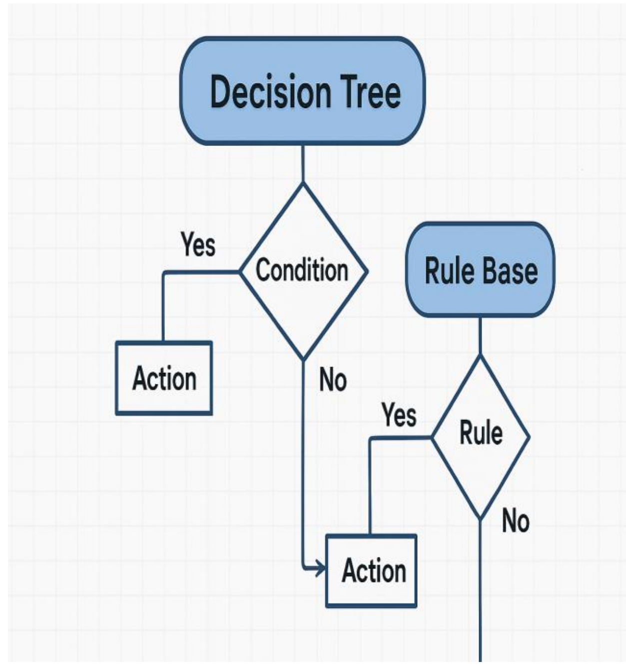


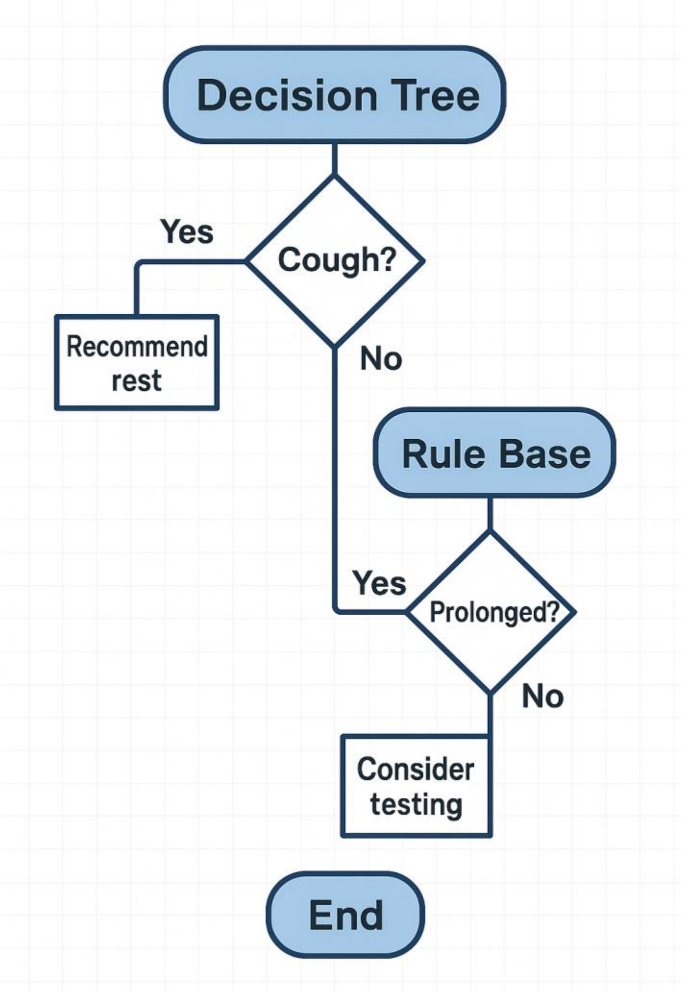
Fig. 1. Hybrid Decision Tree + Rule-Base flowchart

Rule-Based Systems in Medical Diagnosis

Rule-based systems rely on predefined "if–then" rules derived from expert knowledge or empirical data to infer diagnoses. They were foundational in early medical AI, exemplified by systems such as **MYCIN**, which supported infectious disease diagnosis. Rules are typically constructed by domain experts and encode clinical guidelines, protocols, or empirical knowledge. For example:

IF fever AND cough AND shortness of breath
THEN suspect pneumonia

These systems are valued for their transparency, as clinicians can trace the reasoning process behind each diagnostic output. They are particularly effective in domains with well-established diagnostic criteria. However, as the number of rules grows, maintaining and updating the knowledge base becomes increasingly complex. Traditional rule-based systems also



lack adaptability, as they do not learn from new data unless rules are manually revised.

To address these limitations, hybrid approaches that combine rule-based reasoning with machine learning are gaining traction. Such systems leverage the interpretability of rules while incorporating the adaptability of data-driven models, making them more effective in dynamic medical environments (Shortliffe, 1976; Sharma & Rani, 2020; Kononenko, 2001).

Implementation Strategies

Implementing decision trees involves collecting labeled medical datasets—such as electronic health records or clinical trial data—and preprocessing them for model training. Feature selection is guided by clinical relevance, and validation techniques such as cross-validation are employed to ensure robustness. These models can function independently or be embedded within broader Clinical Decision Support Systems (CDSS).

Rule-based systems, by contrast, require collaboration with medical experts to encode diagnostic logic into a structured knowledge base. Inference engines apply these rules to patient data to generate diagnostic outputs. Upgrades are essential to incorporate evolving medical guidelines and emerging clinical knowledge. See Figure 2, Combined decision-tree + rule-based flow chart for cough

Hybrid systems blend both approaches. For example, decision tree-derived rules can be integrated into rule-based frameworks, enabling automated learning while preserving expert oversight. This fusion enhances adaptability, interpretability, and trustworthiness, making AI-driven diagnostic tools more effective in clinical practice.

Fig.2.Combined decision-tree + rule-based flow chart for cough

Case Studies: Diagnosing Diabetes and Pneumonia

Case Study 1: AI-Assisted Diabetes Diagnosis

A hospital introduced an AI-powered diagnostic tool to support early detection of diabetes. The system employed a decision tree algorithm trained on a comprehensive dataset that included patient demographics, medical history, laboratory values (e.g., fasting glucose, HbA1c), and lifestyle indicators.

From this data, the model derived interpretable rules, such as:

IF fasting glucose > 126 mg/dL AND BMI > 30

THEN diagnosis = Diabetes Positive

This if-then rule code was reviewed by endocrinologists to ensure clinical validity before being embedded into a rule-based interface accessible to healthcare providers. When assessing a new patient, the system combined machine-learned rules with expert-defined criteria to offer diagnostic suggestions and recommend further testing or interventions.

This dual-layered approach enhanced diagnostic precision, maintained transparency, and allowed iterative updates as new clinical data and guidelines emerged.

Case Study 2: Pneumonia Detection in Emergency Care

In a high-pressure emergency department setting, an AI-driven system was deployed to expedite pneumonia diagnosis. The decision tree model was trained on clinical indicators

such as presenting symptoms (fever, cough, chest discomfort), vital signs (temperature, respiratory rate, oxygen saturation), laboratory markers (CRP, white blood cell count), and radiology reports.

One extracted decision rule was:

IF temperature > 38°C AND respiratory rate > 22/min AND CRP > 10 mg/L

THEN diagnosis = Pneumonia Likely

These rules were validated by pulmonology experts and integrated with established clinical scoring systems such as CURB-65. Upon patient intake, the system analyzed real-time clinical inputs, estimated pneumonia likelihood, flagged high-risk cases, and suggested next steps such as imaging or antibiotic initiation.

This implementation led to faster diagnostic decisions, reduced missed cases, and supported continuous learning through feedback from confirmed diagnoses.

3.0 Conclusion

Decision tree algorithms and rule-based systems each contribute uniquely to the advancement of AI in medical diagnostics. Decision trees provide automated, data-driven insights, while rule-based systems ensure clinical transparency and adherence to expert knowledge. Their integration enables the creation of diagnostic tools that are not only accurate and scalable but also interpretable and trustworthy.

As healthcare continues to evolve, collaboration between medical professionals and data scientists will be essential to refine these systems and unlock their full potential. Continued research and innovation will drive the development of AI solutions that enhance clinical decision-making, strengthen evidence-based practice, and improve patient outcomes.

4.0 References

Abdulqader, H. A., & Abdulazeez, A. (2024). Review on decision tree algorithm in healthcare applications. *Indonesian Journal of Computer Science*, *13*(3). https://doi.org/10.33022/ijcs.v13i3.4026

Azar, A. T., & El-Metwally, S. M. (2012). undefined. *Neural Computing and Applications*, *23*(7-8), 2387-2403.

https://doi.org/10.1007/s00521-012-1196-7

Baronti, V., & Araujo, E. (2021). undefined. *2021 International Conference on Decision Aid Sciences and Application (DASA)*, 706-711. https://doi.org/10.1109/dasa53625.2021.9682283

Baronti, V., & Araujo, E. (2021). Diagnosing support expert system for symptomatic fever onset of Kawasaki disease. *2021 International Conference on Decision Aid Sciences and Application (DASA)*, 706-711. https://doi.org/10.1109/dasa53625.2021.9682283

Han, J., Kamber, M., & Pei, J. (2012). Advanced pattern mining. *Data Mining*, 279-325. https://doi.org/10.1016/b978-0-12-381479-1.00007-1

Kim, S., Kim, D., Kim, M., Ko, H., & Jeong, O. (2024). XAI-based clinical decision support system: A systematic review. https://doi.org/10.20944/preprints202406.0721.v1

Kononenko, I. (2001). Machine learning for medical diagnosis: History, state of the art and perspective. *Artificial Intelligence in Medicine*, *23*(1), 89-109. https://doi.org/10.1016/s0933-3657(01)00077-x

Kune, M. M., & Panggabean, E. (2020). Diagnosing asthma expert system with case base reasoning methods. *Journal Of Computer Networks, Architecture and High Performance Computing*, *2*(2), 187-190. https://doi.org/10.47709/cnapc.v2i2.398

Lhotska, L., & Vlcek, T. (n.d.). Efficiency enhancement of rule-based expert systems. *Proceedings of 15th IEEE Symposium on Computer-Based Medical Systems (CBMS 2002)*, 53-58. https://doi.org/10.1109/cbms.2002.1011354

Mustafa, E. M., Saad, M. M., & Rizkallah, L. W. (2023). undefined. *Journal of Engineering and Applied Science*, *70*(1). https://doi.org/10.1186/s44147-023-00315-4

Patel, V. L., Shortliffe, E. H., Stefanelli, M., Szolovits, P., Berthold, M. R., Bellazzi, R., & Abu-Hanna, A. (2009). The coming of age of artificial intelligence in medicine. *Artificial Intelligence in Medicine*, *46*(1), 5-17. https://doi.org/10.1016/j.artmed.2008.07.017

Quinlan, J. (1990). Probabilistic decision trees. *Machine Learning*, 140-152. https://doi.org/10.1016/b978-0-08-051055-2.50011-0

Shakhgeldyan, K., Gribova, V., Shalfeeva, E., & Potapenko, B. (2023). Hybrid clinical decision support system in cardiovascular medicine. https://doi.org/10.2139/ssrn.4522020

Shishehchi, S., & Banihashem, S. Y. (2021). undefined. *Smart Health*, *21*, 100192.

https://doi.org/10.1016/j.smhl.2021.100192

Shortliffe, E. H. (1976). Future directions for MYCIN. *Computer-Based Medical Consultations: Mycin*, 205-232. https://doi.org/10.1016/b978-0-444-00179-5.50012-3