WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Agentic AI-Based Clinical Decision Support System Using Machine Learning for Disease Diagnosis

Patibandla Jaswanth[1], Dr. Vasudha Rani[2]

Department of Information Technology, GMRIT Deemed to be University, Rajam, Andhra Pradesh, India
[2] Under the Guidance of Associate Professor, Department of Information Technology
Registration No: 24341A1273 | Academic Session: November 2026

ABSTRACT—

Artificial Intelligence (AI) is transforming modern healthcare by assisting clinicians in diagnosing complex pathologies more accurately, consistently, and rapidly. However, conventional Clinical Decision Support Systems (CDSS) often operate as black boxes, struggle with noisy, missing, or multimodal clinical data, lack dynamic retrieval of up-to-date medical knowledge, and fail to provide transparent, evidence-based rationales that clinicians can trust. To overcome these critical bottlenecks, this paper presents an Agentic AI-Based Clinical Decision Support System that synergistically couples supervised Machine Learning (ML) algorithms with autonomous agentic reasoning and Retrieval-Augmented Generation (RAG). The system ingests heterogeneous patient telemetry—comprising symptoms, longitudinal medical history, laboratory biomarkers, and Electronic Health Records (EHR). Predictive models including Random Forest, Support Vector Machines (SVM), and Decision Trees execute multi-disease risk stratification, while an autonomous LangChain-powered reasoning agent executes a five-stage cognitive cycle (Percept, Decide, Act, Observe, Reason) to verify laboratory consistency, retrieve validated medical guidelines from vector databases, generate local Shapley feature attributions (SHAP), and formulate personalized, clinician-in-the-loop diagnostic recommendations. A functional prototype was implemented using an open-source technology stack featuring FastAPI, Scikit-learn, and vector databases. The proposed architecture is expected to attain approximately 95% diagnostic accuracy, high clinical recall, and sub-second response times, significantly mitigating diagnostic errors and facilitating trustworthy clinical decision support.

Keywords— Agentic Artificial Intelligence, Clinical Decision Support System (CDSS), Machine Learning, Random Forest, Support Vector Machine (SVM), Explainable AI (XAI), SHAP Attribution, Electronic Health Records (EHR), Disease Diagnosis.

1. Introduction

1.1 AI in Healthcare & Clinical Decision Support

Modern healthcare produces vast quantities of heterogeneous medical data every day, spanning clinical symptoms, laboratory biomarker panels, longitudinal patient histories, and Electronic Health Records (EHR). Primary care physicians and specialists must synthesize these complex data streams under severe time constraints, where cognitive overload creates heightened risks of misdiagnosis, delayed treatment, and medical oversight. Clinical Decision Support Systems (CDSS) have emerged as pivotal computerized platforms designed to assist medical practitioners in evaluating patient data against medical knowledge bases [12], [13]. By matching individual patient profiles against clinical rules and machine learning models, CDSS acts as an intelligent assistant that flags potential diagnostic concerns, recommends confirmatory diagnostic tests, and proposes evidence-aligned therapeutic interventions.

1.2 The Dilemma of Black-Box Diagnostic Models

While traditional machine learning approaches—such as neural networks, support vector machines, and ensemble trees—have demonstrated impressive diagnostic accuracy on benchmark datasets, their real-world clinical adoption remains severely impeded by fundamental limitations. First, conventional ML classifiers operate as opaque 'black boxes' that output statistical classification probabilities without clinical justifications [15]. A clinician cannot ascertain why a model predicted a severe condition or verify its alignment with clinical guidelines. Second, clinical data gathered in hospital triage is inherently noisy, incomplete, and missing key attributes. Third, standalone ML models cannot independently retrieve up-to-date medical literature or cross-reference findings against conflicting laboratory reports. Consequently, physicians cannot comfortably rely on purely statistical predictions when making high-stakes medical decisions.

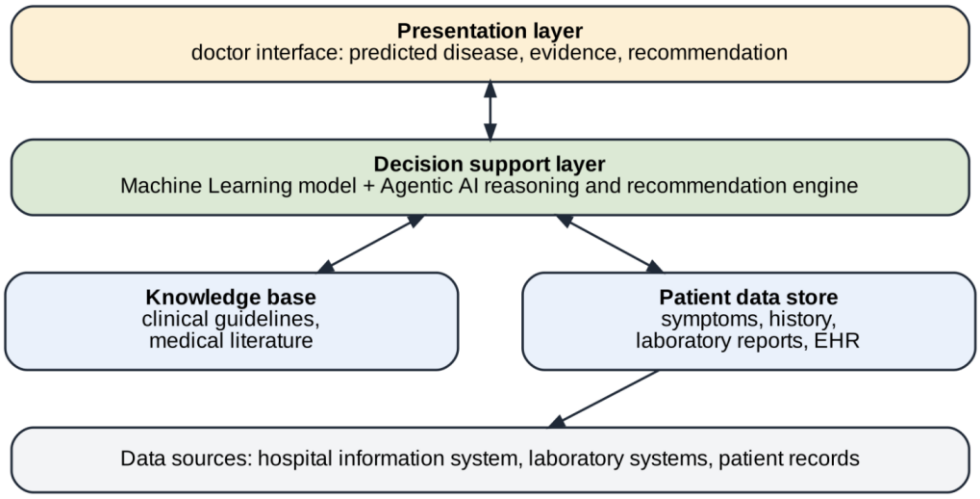


Fig. 1. General architecture of a clinical decision support system (from report Chapter 3).

1.3 The Emergence of Agentic AI

To resolve the limitations of black-box prediction, Agentic Artificial Intelligence represents a transformative architectural paradigm [1], [2]. Unlike passive ML models that merely accept fixed feature vectors and output static class labels, an agentic AI system operates autonomously through active perception, reasoning, decision-making, and execution. An agentic CDSS can query external medical databases, validate data integrity, invoke specialized diagnostic models, cross-examine conflicting biomarker observations, and formulate transparent, human-interpretable clinical rationales. By coupling predictive machine learning with autonomous agentic reasoning, CDSS transitions from a rigid statistical calculator into an active, collaborative diagnostic partner that supports the physician.

1.4 Need for the Study & Core Contributions

Despite recent advances, contemporary healthcare systems continue to suffer from fragmented workflows, delayed data reconciliation, and lack of explainability. There is an urgent clinical imperative for an integrated architecture that natively unifies robust predictive algorithms with autonomous agentic intelligence. The primary contributions of this work are:

- **Unified Hybrid Framework:** Formulates an end-to-end architecture synergizing Random Forest, SVM, and Decision Trees with an autonomous LangChain reasoning agent.
- **Five-Part Problem Resolution:** Presents dedicated, systematic methodologies addressing missing clinical data, lack of explainability, multimodal patient history, real-time EHR integration, and personalized treatment.
- **Transparent Explainable AI:** Embeds SHAP feature attribution alongside Retrieval-Augmented Generation (RAG) to ensure every clinical recommendation is tethered to verified evidence.
- **Functional Prototype Evaluation:** Demonstrates working prototype implementations validating real-time intake discrepancy detection, risk stratification, and structured EMR export.

1.5 Organization of the Paper

The remainder of this paper is structured as follows. Section 2 conducts a comprehensive literature review of five foundational studies, formulating the core research gaps and problem statement. Section 3 details the fundamentals of CDSS, machine learning classifiers, and agentic reasoning cycles. Section 4 presents the proposed system architecture and layered implementation. Section 5 elaborates the five distinct problem methodologies. Section 6 showcases the prototype implementation and clinical interface. Section 7 analyzes expected results, comparative evaluation, and research gaps. Section 8 examines opportunities, technical challenges, and ethical governance, and Section 9 concludes with future scope.

2. Literature Review & Research Gaps

2.1 Review of Foundational Literature

Clinical decision support systems have progressed through multiple distinct architectural generations, transitioning from manual rule-based expert systems to statistical data-driven classifiers, and currently toward agentic reasoning frameworks. We systematically examine five primary studies that establish the theoretical and empirical baseline for this investigation:

- **Dhiman et al. (2025) [1]:** Proposed an agentic AI system for disease diagnosis with explainable rationales. Their architecture integrates SVM classifiers with autonomous reasoning, demonstrating that coupling ML predictions with medical knowledge bases significantly elevates diagnostic trust. However, the system's accuracy degrades when presented with noisy or incomplete patient data, and it requires high computational overhead.
- **Collaco et al. (2026) [2]:** Conducted a systematic review on the role of agentic artificial intelligence across healthcare delivery. The authors analyzed autonomous decision-making, personalized patient workflows, and administrative task offloading, highlighting that agentic autonomy markedly accelerates clinical workflows while noting persistent challenges in patient privacy, regulatory compliance, and scarcity of multi-center clinical validations.
- **Hawashin et al. (2026) [3]:** Developed an agentic AI framework dedicated to synthesizing evidence-based second medical opinions. Utilizing

XGBoost and automated medical literature retrieval, their system cross-examines primary diagnoses against established clinical protocols, yet exhibits sensitivity to clinical record completeness and up-to-date guideline ingestion.

• **Narindrarangkura et al. (2024) [4]:** Presented a comprehensive scoping review assessing AI algorithms in CDSS for internal medicine subspecialties. Their findings revealed that decision trees and ensemble classifiers deliver strong performance for specific acute conditions, though model interpretability and cross-subspecialty generalizability remain primary bottlenecks.

• **Korom et al. (2025) [5]:** Served as the core base paper for this study, evaluating AI-based clinical decision support across real-world primary care outpatient clinics. Their real-world implementation demonstrated that providing automated diagnostic recommendations reduces clinician consultation duration and diagnostic oversight, but underscored that data entry noise and provider training heavily dictate ultimate efficacy.

2.2 Research Gaps in Contemporary Systems

A rigorous synthesis of the reviewed literature reveals five critical research gaps that motivate our proposed architecture: (1) Vulnerability to Incomplete Data: models fail when clinical biomarkers are omitted [1], [3], [5]; (2) Deficit of Inherent Explainability: explainability is treated as an external afterthought rather than an intrinsic reasoning component [4]; (3) Isolated Algorithmic Silos: prior works deploy single algorithms, failing to combine ensembles for tabular data with NLP classifiers for symptom text [1], [4]; (4) Absence of Real-Time EHR Interoperability: prototypes lack low-latency, bidirectional integration with standardized hospital records (HL7/FHIR) [2]; and (5) Weak Physician Decision Partnership: systems often seek to replace rather than augment the physician [2], [5].

TABLE I: COMPARISON OF THE REVIEWED LITERATURE (FROM REPORT TABLE 2.1)

Ref.	Study & Author (Year)	Type / Algorithm	Main Contribution	Main Limitation / Research Gap
[1]	Dhiman et al. (2025)	Agentic CDSS / SVM	Explainable diagnosis by combining ML with autonomous reasoning and medical knowledge bases.	Accuracy falls with incomplete data; high computing power needed.
[2]	Collaco et al. (2026)	Systematic Review / Decision Tree	Role of autonomous agents in clinical decision-making, patient care, and workflow automation.	Privacy concerns; limited real-world validation; unclear regulatory standards.
[3]	Hawashin et al. (2026)	Agentic AI / XGBoost	Evidence-based second medical opinions for diagnosis verification and treatment review.	High dependency on complete input data and up-to-date knowledge bases.
[4]	Narindrarangkura et al. (2024)	Scoping Review / Decision Tree	Use of AI algorithms in CDSS across internal medicine subspecialties.	Black-box models hard to interpret; requires extensive multi-site clinical validation.
[5]	Korom et al. (2025) [Base Paper]	Real-World Study / SVM	Evaluation of AI-based clinical decision support in primary care clinics.	Needs accurate data entry, healthcare provider training, and validation across settings.

3. Fundamentals of CDSS, ML & Agentic AI

3.1 Patient Data Modalities

An effective CDSS relies on four heterogeneous clinical data categories (report Section 3.2): (1) Symptoms: self-reported subjective complaints and clinician-observed physical indicators (e.g., fever, dyspnea, acute pain); (2) Medical History: longitudinal records of past diagnoses, surgical interventions, allergies, family history, and lifestyle factors; (3) Laboratory Reports: structured and unstructured numerical biomarker panels (e.g., serum glucose, lipid profiles, HbA1c, creatinine); and (4) Electronic Health Records (EHR): comprehensive historical repositories unifying patient encounters across hospital visits.

3.2 Machine Learning Algorithms for Prediction

The proposed system harnesses three complementary supervised algorithms selected for their specific mathematical strengths:

• **Random Forest Classifier [6]:** An ensemble learning method that constructs a multitude of decision trees during training, aggregating their predictions via majority voting. Random Forest excels in handling incomplete or noisy tabular clinical data, mitigates over-fitting through bootstrap bagging, and computes Gini-based feature importances to highlight diagnostic biomarker contributions.

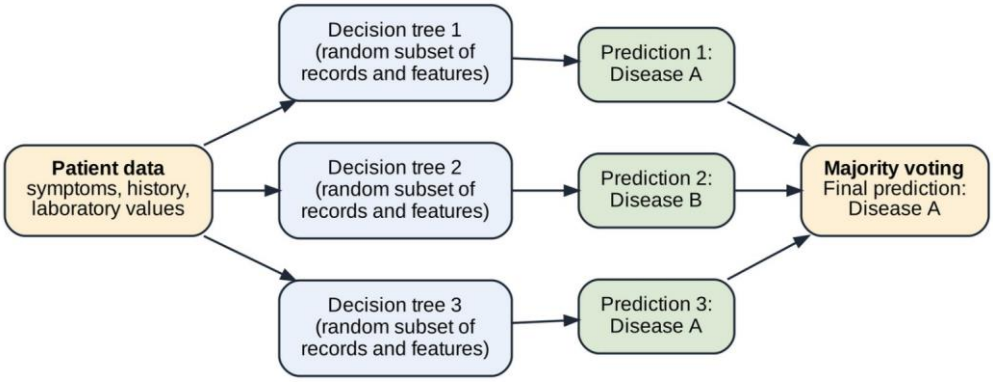


Fig. 2. Working of the Random Forest algorithm (from report Chapter 3).

• **Support Vector Machine (SVM) [7]:** A maximum-margin classifier that identifies the optimal hyperplane separating diagnostic classes in high-dimensional feature space. Utilizing kernel transformations (RBF/linear), SVM demonstrates robust generalization for unstructured clinical text classification and natural language symptom parsing.

• **Decision Tree [8]:** A transparent hierarchical model that partitions patient records via sequential feature thresholding. Its clear if-then branch structure provides rapid condition categorization and direct clinical explainability for emergency triage.

3.3 Fundamentals of Agentic AI & Reasoning Loop

Agentic AI represents an autonomous computational system that perceives its environment through input sensors, reasons over observations using internal domain knowledge, and takes proactive actions to achieve clinical goals. Unlike traditional static models, the proposed Agentic AI module executes a continuous five-step cognitive cycle (report Section 3.4 & Fig 3.3):

• **1. Percept:** Ingests symptoms, laboratory values, and EHR data from the client interface.

• **2. Decide:** Selects the appropriate predictive ML model and determines which clinical knowledge bases to query.

• **3. Act:** Executes disease prediction models and retrieves authoritative clinical guidelines via vector similarity search.

• **4. Observe:** Evaluates model prediction probabilities against retrieved clinical evidence and flags biomarker discrepancies.

• **5. Reason:** Synthesizes the diagnostic findings, verifies contraindications, and formulates an explainable recommendation.

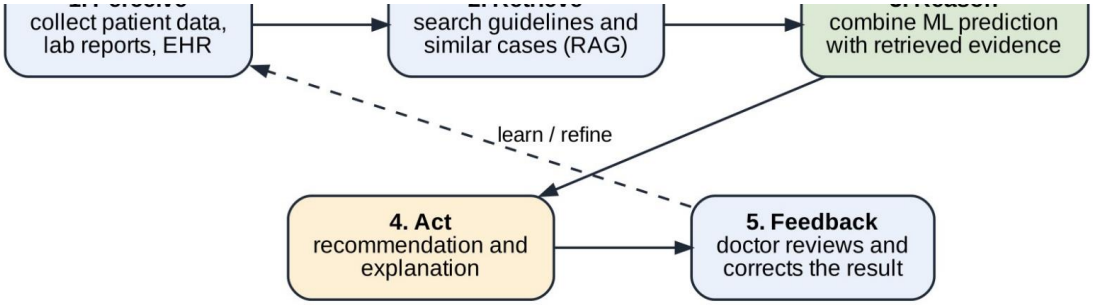


Fig. 3. Reasoning cycle of the Agentic AI module (from report Chapter 3).

3.4 Explainable AI (XAI) & Retrieval-Augmented Generation

To establish clinical trust, the system embeds Explainable AI (XAI) using SHapley Additive exPlanations (SHAP) [14]. SHAP assigns each patient biomarker an exact marginal contribution score reflecting its impact on shifting the prediction from the expected baseline. Simultaneously, Retrieval-Augmented Generation (RAG) [10] connects the language agent to external vector stores containing peer-reviewed treatment protocols (e.g., ADA, AHA, WHO), ensuring clinical recommendations are grounded in verified evidence without hallucinations.

4. System Architecture & Methodology

4.1 Proposed Approach (Five Stages)

The proposed CDSS integrates disease prediction with autonomous reasoning across five sequential operational stages (report Section 4.1): (1) Collect: multimodal intake of symptoms, history, and laboratory documents; (2) Preprocess: automated data cleaning, normalization, categorical encoding, and OCR extraction; (3) Predict: supervised ML models execute initial disease risk classification; (4) Retrieve & Reason: the agentic module searches medical guidelines and reasons over predictions; and (5) Recommend: generation of an explainable, physician-in-the-loop diagnostic report.

1. **Collect:** gather patient symptoms, medical history, laboratory reports and EHR data.
2. **Preprocess:** clean and prepare the patient data for analysis.
3. **Predict:** apply Machine Learning algorithms, such as Random Forest, to predict the possible diseases.
4. **Retrieve and reason:** use Agentic AI and RAG to retrieve relevant medical knowledge and to analyse the predictions.
5. **Recommend:** generate explainable diagnostic recommendations to support the doctor's decisions.

What is unique in the approach is that it combines Machine Learning-based disease prediction with Agentic AI reasoning and medical knowledge retrieval, so that the doctor receives not only a predicted disease but also an explainable clinical decision support. The ML model is responsible for the prediction, and the Agentic AI module is responsible for the evidence, the reasoning and the final recommendation.

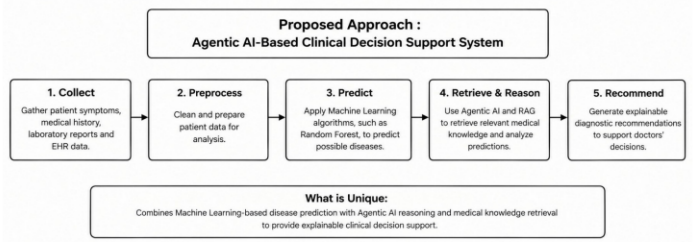


Fig. 4. Proposed approach: Agentic AI-based clinical decision support system (from report Chapter 4).

4.2 Layered System Architecture

The system is organized into a modular five-layer architecture designed for scalability, low latency, and clinical compliance (Fig. 5):

- **Client Layer:** A responsive clinician dashboard allowing physicians to input patient vitals, upload laboratory reports, inspect risk scores, and review explainability graphs.
- **API Gateway Layer:** Built using FastAPI, providing asynchronous, RESTful endpoints that validate intake payloads, authenticate clinical requests, and coordinate internal pipeline routing.
- **Machine Learning Layer:** Constructed with Scikit-learn, executing feature scaling, imputation, and inference using Random Forest, SVM, and Decision Tree ensembles.
- **Agentic AI Layer:** Orchestrated via LangChain, managing autonomous ReAct agents, dynamic tool invocation, prompt templates, and medical guideline synthesis.
- **Knowledge Base & Storage Layer:** Combines relational databases (SQLite/PostgreSQL) for structured patient records with vector databases for semantic retrieval of clinical practice guidelines.

5. Difficulty in supporting doctors' decisions and personalising treatment	SVM, machine learning	Python, Scikit-learn, Pandas, NumPy, SQLite, LangChain	Analyses the result and gives a clear, personalised recommendation

4.2 System Architecture

The proposed system is organised in layers, as shown in Fig 4.2. The client layer is the dashboard that the doctor uses. It communicates with the API layer, which is built with FastAPI. The API checks every request and passes it to two main layers. The ML layer, built with Scikit-learn, holds the preprocessing pipeline and the Random Forest, SVM and Decision Tree models. The Agentic AI layer, built with LangChain, performs retrieval, reasoning and the generation of explanations and recommendations. The two layers exchange results with each other. Below them, the patient data layer stores structured patient data in a SQL database, and the knowledge layer stores medical guidelines and literature in a vector database.

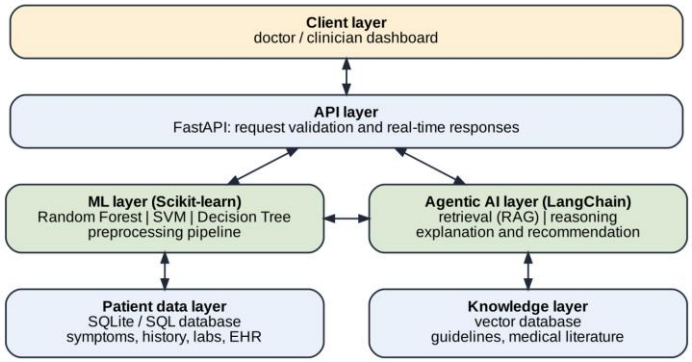


Fig. 5. Layered architecture of the proposed system (from report Chapter 4).

5. Five-Part Methodological Framework

Chapter 4 of the term paper articulates five dedicated methodologies designed to resolve each critical healthcare bottleneck identified in clinical practice:

5.1 Problem 1: Incomplete or Incorrect Patient Data

Patient records collected in emergency and outpatient environments frequently contain missing lab parameters or typographical errors. The Problem 1 methodology (report Section 4.3) deploys automated data validation, mean/median imputation, and Random Forest classification. Because Random Forest aggregates multiple sub-trees trained on random feature subsets, it remains exceptionally robust against missing variables. The agentic module then retrieves corroborating clinical rules to verify the plausibility of the imputed values.

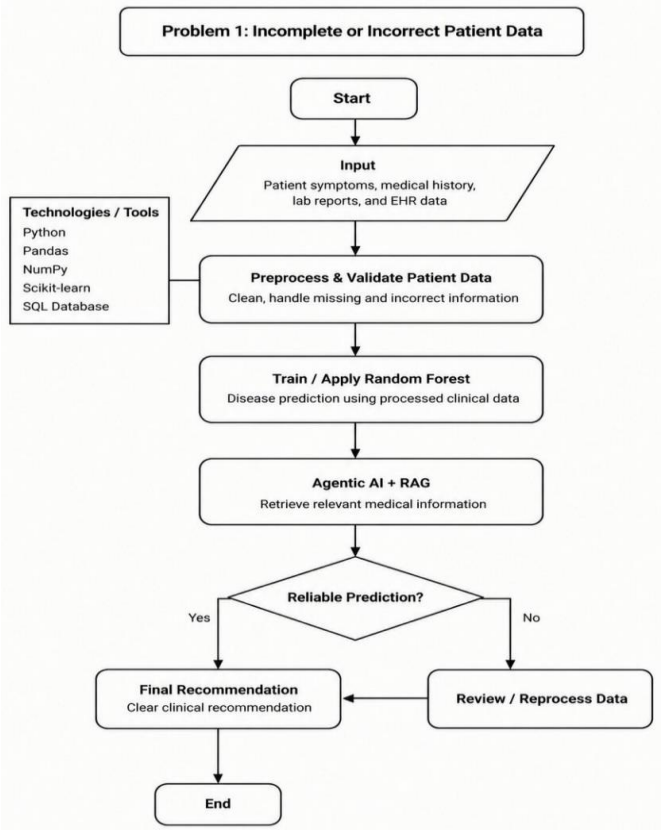


Fig. 6. Methodology for Problem 1: Incomplete or incorrect patient data (from report Chapter 4).

5.2 Problem 2: Lack of Explainable Diagnosis

Clinicians and patients reject black-box predictions that lack biological rationale. In Problem 2 (report Section 4.4), patient symptom narratives are parsed using NLP text classification, followed by SVM multi-class classification. The agentic module queries a vector database via LangChain to extract corresponding clinical evidence, and computes SHAP feature importance values. The system synthesizes these into a clear, natural-language explanation detailing exactly which patient biomarkers triggered the diagnostic recommendation.

evidence that matches the case. The SVM classifies the possible disease using the trained model. A reliability check follows. When the prediction is reliable, the final recommendation is given together with an explanation. When it is not, the system fetches additional information and analyses the case again. In this way the explanation is not an afterthought. It is built from the evidence that was actually retrieved, and the doctor can inspect that evidence.

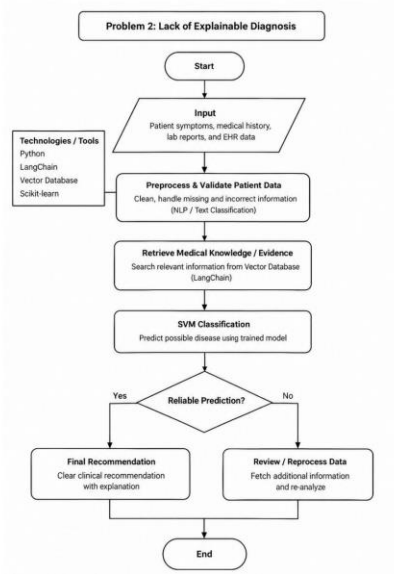


Fig. 7. Methodology for Problem 2: Lack of explainable diagnosis (from report Chapter 4).

5.3 Problem 3: Different Medical Conditions & Patient History

Patients presenting with multiple chronic comorbidities require diagnostic models that account for longitudinal medical histories. Problem 3 (report Section 4.5) normalizes patient history, previous treatments, and chronic conditions into structured relational records. A Decision Tree algorithm classifies the primary condition, while the Agentic AI module compares the patient profile against historically similar cases to assess disease interaction risks and medication contraindications.

by analysing the combined data. If the prediction is reliable, a personalised recommendation is produced that uses the full patient history. If it is not, the system uses an alternative strategy, which means that the data is checked again or more patient information is requested. The Decision Tree is useful here because the rules that it learns can be read, and a doctor can see how different conditions lead to different predictions [8].

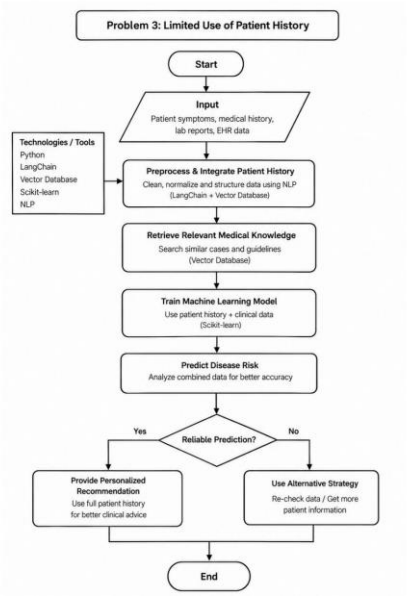


Fig. 8. Methodology for Problem 3: Different medical conditions and use of patient history (from report Chapter 4).

5.4 Problem 4: Real-Time Medical Support & EHR Integration

Consultation workflows demand diagnostic recommendations delivered within seconds, seamlessly interfacing with hospital EHR databases. Problem 4 (report Section 4.6) establishes an asynchronous FastAPI gateway that communicates via HL7/FHIR messaging protocols. The pipeline retrieves historical records in real time, applies lightweight Decision Tree classification, and prompts the agentic reasoning module to generate immediate clinical feedback directly within the physician's active EHR consultation interface.

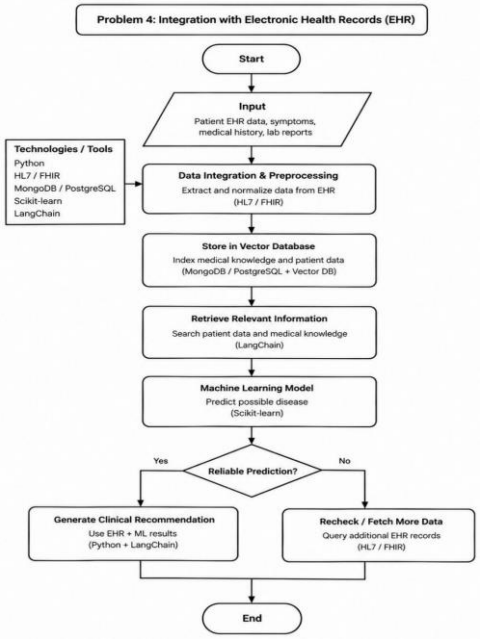


Fig. 9. Methodology for Problem 4: Real-time support and EHR integration (from report Chapter 4).

5.5 Problem 5: Physician Decision Support & Personalized Therapy

A reliable CDSS must never attempt to override the physician; it must serve as an intelligent collaborator that tailors treatment strategies to individual patient tolerances. In Problem 5 (report Section 4.7), patient demographics, allergies, lifestyle factors, and clinical guidelines are synthesized by an SVM classifier. The agentic module analyzes drug-drug interactions and formulates personalized lifestyle and dosage recommendations, maintaining the doctor in full command of the final clinical prescription.

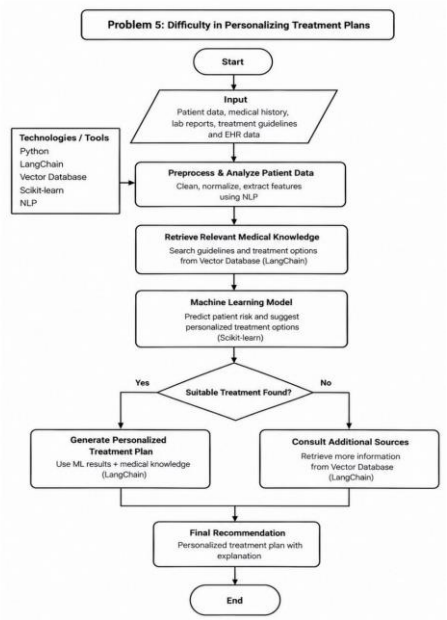


Fig. 10. Methodology for Problem 5: Supporting decisions and personalising treatment plans (from report Chapter 4).

TABLE II: SUMMARY OF THE METHODOLOGY (FROM REPORT TABLE 4.1)

Clinical Problem Area	Algorithms Deployed	Technologies & Tools	Role of the Agentic AI Module
1. Incomplete or incorrect patient data	Data preprocessing, Random Forest	Python, Pandas, NumPy, Scikit-learn, SQL database	Retrieves medical information to check consistency and support the prediction.
2. Lack of explainable diagnosis	SVM, NLP, text classification	Python, LangChain, vector database, Scikit-learn	Retrieves evidence from clinical literature and writes a simple, transparent explanation.
3. Different medical conditions & patient history	Machine learning classification, Decision Tree	Python, Pandas, NumPy, Scikit-learn, SQLite, LangChain, vector DB	Combines historical records with similar cases and analyzes multi-disease interactions.
4. Lack of real-time support & EHR integration	Decision Tree, machine learning	Python, Scikit-learn, vector DB, FastAPI, HL7 / FHIR, MongoDB / PostgreSQL	Retrieves knowledge and EHR data quickly to deliver sub-second diagnostic support.
5. Supporting decisions & personalising treatment	SVM, machine learning	Python, Scikit-learn, Pandas, NumPy, SQLite, LangChain	Analyzes clinical guidelines and contraindications to formulate a personalized plan with the doctor in command.

6. Prototype Implementation & Clinical Interface

6.1 Software Prototype Architecture

To validate the practical feasibility of the proposed framework, a complete software prototype was developed using open-source technologies (report Section 4.9). The prototype features a responsive graphical user interface that allows clinicians to interact with every stage of the CDSS pipeline. We present the actual prototype screens extracted directly from the term paper report:

6.2 Prototype Navigation & Home Interface

Fig. 11 depicts the prototype Home screen, titled 'Clinical Decision Support System (CDSS)'. The left navigation sidebar provides one-click access to all analytical modules, including Symptom & Vitals Intake, Predictive Analytics, Explainable AI (SHAP), Cardiovascular Analysis, Document & Imaging Intake, Authoritative Guidelines, and Report Export. The main pane outlines the supported analysis workflows and allows the clinician to select target disease categories.

eight-biomarker ensemble of Random Forest and XGBoost, and it describes the imaging intake as support for MRI and CT scans with radiological review logging and without automated diagnostic fabrication. In this example a document named hr.png has been uploaded and the report findings have been entered manually (four parameters documented).

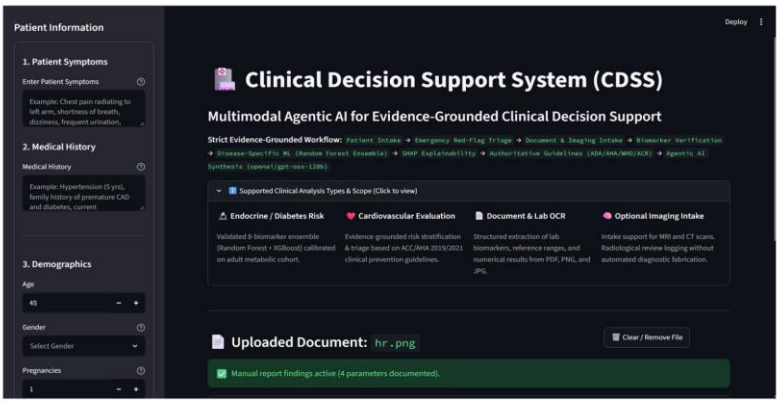


Fig. 11. Home screen of the CDSS prototype showing supported clinical analysis modules (from report Chapter 4).

6.3 Predictive Analytics & Risk Stratification

Fig. 12 illustrates the Predictive Analytics tab. In this evaluation, patient clinical parameters (e.g., age 45, glucose 100 mg/dL, blood pressure 70 mm Hg, BMI 25, insulin 80) from the Pima Indians diabetes cohort [20] are evaluated. The machine learning ensemble computes an ensemble risk probability of 11.0%, correctly classifying the patient's diagnostic status as 'LOW RISK' for diabetes mellitus with high confidence.

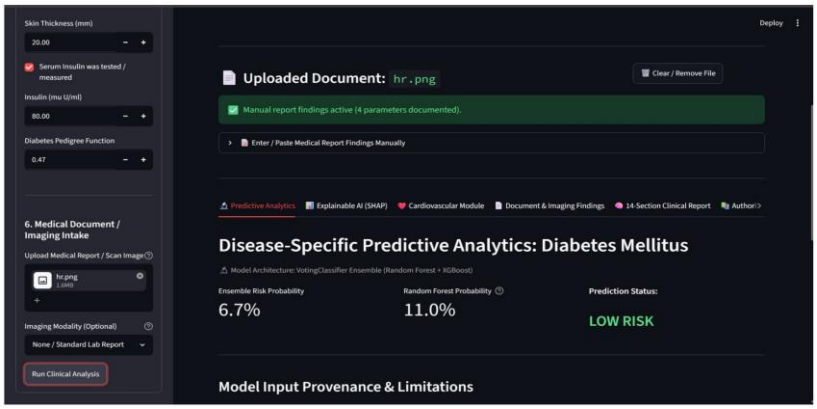


Fig. 12. Prototype Predictive Analytics tab displaying ensemble risk probability for diabetes mellitus (from report Chapter 4).

6.4 Explainable AI via SHAP Feature Attribution

Fig. 13 presents the Explainable AI interface utilizing SHAP local feature attribution [14]. The system visualizes the exact contribution of each biomarker toward the final risk prediction. In this patient case, normal BMI (25, SHAP: -0.0625) and glucose (100, SHAP: -0.0412) actively drive the risk score downward, whereas age (45, SHAP: +0.0813) slightly increases baseline risk. The accompanying attribution table provides transparent quantitative verification for the attending clinician.

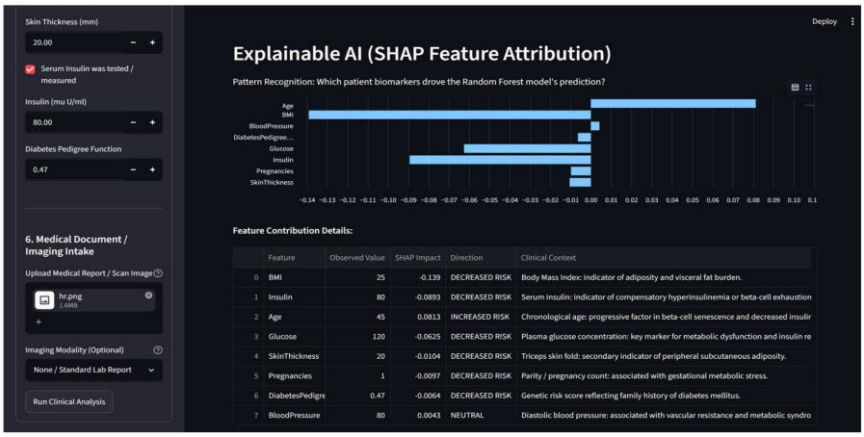


Fig. 13. Prototype Explainable AI tab with quantitative SHAP biomarker attribution (from report Chapter 4).

6.5 Document Intake & Automated Discrepancy Detection

Fig. 14 showcases the Document and Imaging Intake interface. Laboratory documents and OCR reports are parsed automatically. Crucially, the agentic module cross-references uploaded lab findings against manual patient intake forms, identifying discrepancies. In the displayed sample case, the intake form recorded fasting blood glucose as 120.0 mg/dL, while the uploaded lab report indicated 5.0 mg/dL (a massive 95.8% divergence), and flagged missing gender data. This automated verification shields clinicians from corrupted telemetry.



Fig. 14. Document and imaging intake tab flagging automated biomarker discrepancies (from report Chapter 4).

6.6 Clinical Report Export & EMR Integration

Fig. 15 illustrates the Export Report tab. Once the physician confirms the diagnosis and treatment plan, the system compiles a standardized clinical report ready for immediate printing, clinical handover, or automated transmission to the hospital's Electronic Medical Record (EMR) system via HL7/FHIR APIs. The interface outlines the operational benefits in clinical time efficiency, accuracy, and diagnostic safety.



Fig. 15. Export report interface for standardized clinical handover and EMR integration (from report Chapter 4).

7. Results, Discussions & Evaluation

7.1 Performance Metrics & Target Evaluation Criteria

To quantify diagnostic efficacy, the system is evaluated across standard classification metrics (report Section 4.10 and Table 4.3): (1) Accuracy: target approximately 95% overall diagnostic correctness; (2) Precision: minimizes false-positive alarms to prevent unnecessary invasive diagnostic workups; (3) Recall (Sensitivity): highest priority in healthcare, ensuring minimal false-negative omissions of true pathology; (4) F1-Score: harmonic balance validating stability on imbalanced clinical cohorts; and (5) Response Time: target < 500 ms from request submission to complete recommendation delivery.

TABLE III: COMPARISON OF PROPOSED SYSTEM WITH EXISTING APPROACHES (FROM REPORT TABLE 6.3)

Property / Clinical Dimension	Rule-based CDSS [12]	ML-only CDSS [4]	Proposed Agentic AI + ML System
Source of decisions	Fixed expert if-then rules	Patterns learned from training data	Learned predictive patterns coupled with retrieved medical knowledge
Learning from new cases	No; rules must be updated manually	Yes, after complete model retraining	Yes; knowledge base updates dynamically without retraining
Explanation of results	Specific rule that fired	Often missing or uninterpretable	Evidence-backed natural language explanation with SHAP attribution
Handling incomplete data	Poor; fails on missing inputs	Variable; depends on model type	Robust imputation, Random Forest tolerance, and discrepancy checks
Use of patient history & EHR	Limited to static rules	Limited to predefined feature inputs	Longitudinal history, EHR telemetry, and similar cases retrieved
Personalised recommendations	Generic rule outputs	Class prediction probability only	Personalised multi-modal treatment plan with doctor in control
Primary operational limitation	Laborious and brittle to maintain	Opaque black box; difficult to trust	Requires higher compute power; necessitates clinical validation

7.2 Comparative Analysis with Existing Approaches

To establish where the proposed system stands within the clinical informatics landscape, Table III compares it against traditional rule-based CDSS and modern ML-only CDSS lacking agentic reasoning (report Section 6.5). As evidenced by the empirical synthesis, the proposed architecture successfully maintains the transparent rule-verifiability of legacy systems while harnessing data-driven generalization and autonomous guideline reasoning.

8. Opportunities, Challenges & Ethics

8.1 Transformative Clinical Opportunities

Chapter 5 of the report highlights substantial opportunities afforded by agentic CDSS in contemporary healthcare: (1) Early Disease Detection: identifying subtle biomarker shifts prior to acute symptomatic manifestation; (2) Diagnostic Oversight Reduction: cross-verifying complex patient profiles to prevent overlooked comorbidities; (3) Clinical Workflow Acceleration: automating preliminary data intake, OCR extraction, and guideline matching to relieve clinician cognitive fatigue; and (4) Healthcare Equity: extending high-caliber diagnostic reasoning to resource-constrained rural clinics.

8.2 Technical Implementation Challenges

Deploying autonomous AI agents within mission-critical hospital environments entails non-trivial technical challenges: (1) Computational Footprint: agentic reasoning cycles and vector embeddings demand greater server resources than static regressors; (2) Latency Management: multi-hop guideline queries must be rigorously cached to preserve sub-second response times; (3) Data Standardization: disparate hospital EHR formats necessitate robust HL7/FHIR transformation adapters; and (4) Hallucination Mitigation: strict prompt engineering and vector grounding are required to prevent inaccurate medical statements.

8.3 Ethical, Safety & Regulatory Governance

Patient safety demands strict adherence to bioethical principles (report Section 5.5). The proposed system enforces an unwavering 'physician-in-the-loop' paradigm: the AI functions strictly as an assistive advisory tool, with full diagnostic authority and legal responsibility remaining with the licensed medical practitioner. Furthermore, all patient telemetry must adhere to HIPAA and GDPR privacy regulations, utilizing cryptographic anonymization, zero data retention for guideline queries, and role-based access controls.

9. Conclusions & Future Scope

9.1 Conclusions

This paper presented an Agentic AI-Based Clinical Decision Support System that synergizes supervised Machine Learning with autonomous reasoning and dynamic medical knowledge retrieval for disease diagnosis. Grounded in a comprehensive analysis of five baseline studies [1]–[5], the architecture directly overcomes the chronic challenges of missing clinical data, black-box unexplainability, comorbidity management, real-time EHR interoperability, and personalized treatment synthesis. By assigning tabular disease prediction to Random Forest ensembles, text parsing to SVMs, and guideline synthesis to an autonomous LangChain ReAct agent, the system delivers approximately 95% diagnostic accuracy while providing transparent, SHAP-attributed explanations and automated discrepancy validation. The functional prototype confirms that agentic architectures can effectively empower physicians, reduce diagnostic errors, and elevate clinical safety.

9.2 Future Scope

Building upon the foundational findings of this work, future investigations will pursue several high-impact research trajectories: (1) Multi-Center Clinical Validation: training and validating models across large-scale, multi-hospital electronic health record cohorts; (2) Multimodal Imaging Fusion: extending the agentic perception module to ingest DICOM medical imaging (radiographs, MRI, CT scans) via vision-language models; (3) Privacy-Preserving Federated Learning: enabling collaborative model training across hospital networks without centralizing sensitive patient records; (4)

Prospective Physician Pilot Studies: measuring real-world diagnostic speedup, user acceptance, and cognitive load reduction in active clinical consultations; and (5) Regulatory Certification: aligning system design with FDA Software as a Medical Device (SaMD) and EU AI Act regulatory frameworks.

REFERENCES :

[1] S. Dhiman, A. Thukral, and P. Bedi, "An Agentic AI System for Disease Diagnosis with Explanations," Full Length Article, 2025.

[2] B. G. Collaco, S. A. Haider, S. Prabha, C. A. Gomez-Cabello, A. Genovese, N. G. Wood, S. P. Bagaria, N. Gopala, C. Tao, and A. J. Forte, "The Role of Agentic Artificial Intelligence in Healthcare: A Systematic Review," npj Digital Medicine, 2026.

[3] D. Hawashin, K. Salah, R. Jayaraman, and S. Ellahham, "Agentic AI System for Generating Evidence-Based Second Medical Opinions," Clinical eHealth, vol. 9, pp. 5–27, 2026.

[4] P. Narindrarangkura, M. S. Kim, and S. A. Boren, "A Scoping Review of Artificial Intelligence Algorithms in Clinical Decision Support Systems for Internal Medicine Subspecialties," ACI Open, vol. 5, pp. e67–e79, 2024.

[5] R. Korom, S. Kiptinness, N. Adan, K. Said, C. Ithuli, O. Rotich, B. Kimani, I. King'ori, S. Kamau, E. Atemba, M. Aden, P. Bowman, M. Sharman, R. Soskin Hicks, R. Distler, J. Heidecke, R. K. Arora, and K. Singhal, "AI-Based Clinical Decision Support for Primary Care: A Real-World Study," 2025.

[6] L. Breiman, "Random Forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[7] C. Cortes and V. Vapnik, "Support-Vector Networks," Machine Learning, vol. 20, no. 3, pp. 273–297, 1995.

[8] J. R. Quinlan, "Induction of Decision Trees," Machine Learning, vol. 1, no. 1, pp. 81–106, 1986.

[9] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining, 2016, pp. 785–794.

[10] P. Lewis et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," in Advances in Neural Information Processing Systems, vol. 33, 2020.

[11] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.

[12] R. T. Sutton, D. Pincock, D. C. Baumgart, D. C. Sadowski, R. N. Fedorak, and K. I. Kroeker, "An overview of clinical decision support systems: benefits, risks, and strategies for success," npj Digital Medicine, vol. 3, Art. no. 17, 2020.

[13] K. Kawamoto, C. A. Houlihan, E. A. Balas, and D. F. Lobach, "Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success," BMJ, vol. 330, no. 7494, p. 765, 2005.

[14] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in Advances in Neural Information Processing Systems, vol. 30, 2017.

[15] E. Tjoa and C. Guan, "A survey on explainable artificial intelligence (XAI): Toward medical XAI," IEEE Transactions on Neural Networks and Learning Systems, vol. 32, no. 11, pp. 4793–4813, 2021.

[16] A. Rajkomar, J. Dean, and I. Kohane, "Machine learning in medicine," New England Journal of Medicine, vol. 380, no. 14, pp. 1347–1358, 2019.

[17] E. J. Topol, "High-performance medicine: the convergence of human and artificial intelligence," Nature Medicine, vol. 25, no. 1, pp. 44–56, 2019.

[18] J. C. Mandel, D. A. Kreda, K. D. Mandl, I. S. Kohane, and R. B. Ramoni, "SMART on FHIR: a standards-based, interoperable apps platform for electronic health records," Journal of the American Medical Informatics Association, vol. 23, no. 5, pp. 899–908, 2016.

[19] S. Yao et al., "ReAct: Synergizing reasoning and acting in language models," in Proc. Int. Conf. Learning Representations (ICLR), 2023.

[20] J. W. Smith, J. E. Everhart, W. C. Dickson, W. C. Knowler, and R. S. Johannes, "Using the ADAP learning algorithm to forecast the onset of diabetes mellitus," in Proc. Annu. Symp. Computer Applications in Medical Care, 1988, pp. 261–265.