APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN PREDICTIVE HEALTHCARE SYSTEMS

A Thesis
Submitted towards the Requirement for the Award of Degree of

Doctor of Philosophy

In

COMPUTER SCIENCE & ENGINEERING
Under the Faculty of ENGINEERING & TECHNOLOGY

By

SUGARWAR KALYANI SUDHAKAR
(Enrolment no. 161588517593)

Under the Supervision of

DR. SANTANU SIKDAR
Professor

P.K. UNIVERSITY

2025

P.K. University
NH-27, Vill. Thanra, (P.O. Dinara)
Shivpuri, M.P. India– 473665

www.pkuniversity.edu.in

Supervisor Signature

APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN PREDICTIVE HEALTHCARE SYSTEMS

A Thesis
Submitted towards the Requirement for the Award of Degree of

Doctor of Philosophy

In

COMPUTER SCIENCE & ENGINEERING
Under the Faculty of ENGINEERING & TECHNOLOGY

By

SUGARWAR KALYANI SUDHAKAR
(Enrolment no. 161588517593)

Under the Supervision of

DR. SANTANU SIKDAR
Professor

P.K. UNIVERSITY

2025

P.K. University
NH-27, Vill. Thanra, (P.O. Dinara)
Shivpuri, M.P. India– 473665

www.pkuniversity.edu.in

CERTIFICATE OF THE SUPERVISOR

This is to certify that the work entitled **"Applications of Artificial Intelligence in Predictive Healthcare Systems**" is a piece of research Work done by **Ms. Sugarwar Kalyani Sudhakar** Under My Guidance and Supervision for the degree of Doctor of Philosophy in **Computer Science & Engineering,** Faculty of Engineering and Technology, P.K. University (M.P), India.

I certify that the candidate has put an attendance of more than 240 days with me. To the best of my knowledge and belief the thesis:

I. Embodies the work of the candidate himself/herself.
II. Has duly been completed.
III. Fulfill the requirement of the ordinance relating to the Ph.D. degree of the University.

Signature of the Supervisor

Date:……..

DECLARATION BY THE CANDIDATE

I declare that the thesis entitled "**Applications of Artificial Intelligence in Predictive Healthcare Systems"** is my own work conducted under the supervision of **Dr. Santanu Sikdar** at P.K. University (M.P) India Approved by the Research Degree Committee. I have put in more than 240 days of attendance with the Supervisor at the center.

I further declare that to the best of my knowledge, the thesis does not contain my part of any work that has been submitted for the award of any degree either in this University or in any other University Without proper citation.

Signature of the candidate

Date:……

Place:……

FORWARDING LETTER OF HEAD OF INSTITUTION

The Ph.D. thesis entitled "**Applications of Artificial Intelligence in Predictive Healthcare Systems"** Submitted by **Ms. Sugarwar Kalyani Sudhakar** is forwarded to the University in six copies. The candidate has paid the necessary fees and there are no dues outstanding against him/her.

Name: Seal…….

Date: ……

Place…….....

(Signature of Head of Institution where the Candidate was registered for Ph.D. degree)

Signature of the Supervisor

Date: …….. **Address:……**

Place:……..

ACKNOWLEDGEMENT

The Acknowledgement of a research endeavor is an opportunity to express appreciation for the joint effort of those who have been sources of motivation and support, and who have contributed invaluable knowledge towards achieving success in the pursuit of knowledge.

First, for most, I would like to thank almighty god for the perseverance, courage, and strength bestowed on me to undertake this project.

For most, I extend my deepest appreciation to my esteemed Supervisor, **Prof. (Dr.) Santanu Sikdar,** Professor of Computer Science & Engineering, P.K University Shivpuri (M.P) without whom this work would not have come to fruition. His exceptional guidance, unwavering support kind cooperation, and provision of a conducive work environment have been invaluable beyond measure. I remain indebted to her for instilling in me a relentless pursuit of excellence, a strong sense of honesty, and respect for ethical principles that govern our profession.

Furthermore, I extend my sincere thanks to the **Prof. (Dr.) Shri Jagdeesh Prasad Sharma** Chancellor, P.K University Shivpuri (M.P) for providing the necessary infrastructure and facilities that were essential during my research work. My profound sense of gratitude, faith, and awe also goes to **Prof. (Dr.) Yogesh Chandra Dubey** Vice Chancellor, P.K University Shivpuri (M.P) for his generous and valuable support throughout my research.

I take this pleasant opportunity to express my deep sense of gratitude and reverence to his holiness **Prof. (Dr.) Jitendra Kumar Mishra** Director; P.K University Shivpuri (M.P) for his constant encouragement and for providing me with the required facilities that enabled me to complete my project work.

I would also like to extend my gratitude to and deep appreciation to the respected **Prof (Dr). Bhaskar Nalla,** Dean of Research of P.K University M.P, I/c Ph.D. cell of P.K University Shivpuri (M.P), I also extend my deepest appreciation to **Prof. (Dr.) Jitendra Kumar Malik,** Dean of Faculties, P.K University Shivpuri (M.P) for his valuable support and encouragement throughout my research journey. **Prof (Dr.) Deepesh Naamdev** Registrar of P.K University Shivpuri (M.P) has inspired me with

his encouraging guidance and support right from the beginning of this work despite his busy schedule. I am short of words to thank him for his affectionate behavior and patience throughout the duration of the research work.

I express my gratitude to **Prof. Dr. Aiman Fatima,** Dean of Academics of P.K University Shivpuri M.P.

I owe a great deal of application and gratitude to the entire faculty members Department of Computer Science & Engineering P.K University, Head of Department, Computer Science & Engineering, Prof. (Dr.) Balveer Singh, all teaching and non-teaching staff of the department without their support it would have been difficult to shape my research work

I also express my gratitude to **Miss Nisha Yadav,** Librarian of P.K University along with **Mr. Anand Kumar**, Assistant Librarian, and other staff of P.K University. I am also thankful to **Mr. Alok Khare**, I.T. Cell of P.K. University.

I am deeply grateful to my beloved parents, brothers, sister, and my husband, **Mr. Milind Salve**, for their unwavering love, support, and constant encouragement throughout the course of my research work.

I also wish to express my appreciation and heartfelt thanks to my classmates for their kind cooperation in the smooth running of my research work. Last but not least I would like to honestly appreciate the fact that it is my pleasure to have an opportunity to pursue my higher education at **P.K University, Shivpuri (M.P)**. I owe my permanent debt to it; this is having a special place in my heart for shaping my care.

I must also express my love and affection to my son **Ansh M. Salve** for his understanding and unwavering love during my research tenure, Lastly, I am highly grateful to God, the almighty, for giving me courage and strength to face the most difficult times of life and for especially answering my prayers. I remain grateful for all that I have achieved in my life thus far.

Date: **(Sugarwar Kalyani Sudhakar)**

Place:

LIST OF ABBREVIATIONS

Abbreviation	Full Form
AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
IoT	Internet of Things
ANN	Artificial Neural Network
RBF	Radial Basis Function
SVM	Support Vector Machine
PSO	Particle Swarm Optimization
MFA	Modified Firefly Algorithm
ECG	Electrocardiogram
EMR	Electronic Medical Record
EHR	Electronic Health Record
NLP	Natural Language Processing
CNN	Convolutional Neural Network
KNN	K-Nearest Neighbors
TP	True Positive

TN	True Negative
FP	False Positive
FN	False Negative
ROC	Receiver Operating Characteristic
AUC	Area Under Curve
RMSE	Root Mean Square Error
MAE	Mean Absolute Error
HRV	Heart Rate Variability
RF	Random Forest
RNN	Recurrent Neural Network
LSTM	Long Short-Term Memory
FPR	False Positive Rate
TPR	True Positive Rate
GWO	Grey Wolf Optimizer
GA	Genetic Algorithm
XAI	Explainable Artificial Intelligence
CDSS	Clinical Decision Support System

LIST OF TABLE

LIST OF FIGURES

COPYRIGHT TRANSFER CERTIFICATE

Title of the Thesis: **"Applications of Artificial Intelligence in Predictive Healthcare Systems"**

Candidate's Name: **Ms. Sugarwar Kalyani Sudhakar**

COPY RIGHT TRANSFER

The undersigned hereby assigns to the P.K. University, Shivpuri all copyrights that exist in and for the above thesis submitted for the award of the Ph.D. degree.

Date: **Sugarwar Kalyani Sudhakar**

ABSTRACT

India faces a growing number of cardiovascular diseases patients, with over 30 million currently affected. Diagnosing heart disease is a complex task that requires careful analysis and understanding of patients through regular check-ups. Early detection can help patients take precautions and take regulatory measures. The healthcare industry generates vast amounts of data about patients, making machine learning techniques crucial for analyzing this data.

A previous heart disease diagnosis system relied on Interval Type-2 Fuzzy Logic System (IT2FLS), but it had poor recognition accuracy and training time. This research proposes an efficient heart disease prediction system using modified firefly algorithm based radial basis function with support vector machine (MFA and RBF-SVM). The dataset includes three types of qualities: input, key, and prediction characteristics. Standardization is performed using the min-max standardization approach, followed by a heuristic approach called MFA to manage large amounts of high-lights and extensive records. PCA is used to remove highlights, and RBF-SVM is used to classify highlights as ordinary or heart illnesses.

A PSO algorithm and RBF-based Transductive Support Vector Machines (TSVM) approach are proposed to intelligently and efficiently predict heart disease. After normalization, rough sets based attribute reduction using the PSO algorithm is introduced to find optimal reduction. Finally, classification is performed using RBF-TSVM to predict heart diseases. An Opposition Based Crow Search Optimization (OCSO) technique is applied for attribute reduction followed by RBF-TSVM approach.

Metric values are found based on True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). Experimental outcomes show that the proposed techniques achieve superior performance in terms of accuracy, sensitivity, and specificity compared to existing methods.As heart disease is the leading cause of death globally, predicting cardiac disease is a complex task that requires accurate models. Techniques such as Internet of Things (IoT), cloud computing, machine learning, and deep learning techniques are used to build accurate models. Web-based healthcare systems improve the quality of medical diagnostic decisions, and physicians adopt predictive modeling processes to anticipate clinical risk factors

TABLE OF CONTENTS

CHAPTER-1

INTRODUCTION

1.1 OVERVIEW

In today's information era, several industries are constantly producing vast quantities of data. The rapid growth of AI across many fields has been supported by improvements in hardware that enable the processing of this data and the extraction of significant insights. In healthcare, the rising demand for medical services and the strain on healthcare resources due to a growing global population have contributed to the increasing adoption of AI-based systems. Applications such as diagnostic imaging, individualized treatment planning, and disease prediction and prevention have emerged as part of this transformation. These developments collectively form the foundation of "smart healthcare," in which AI-driven predictive analytics play a central role.

Smart healthcare differs from traditional healthcare by shifting the focus from specialists to patients. By integrating advanced intelligent technologies, smart healthcare aims to build a system centered on patient needs, experiences, and participation. This paradigm emphasizes data processing, knowledge discovery from structured and unstructured information, cross-domain insights, and improved decision-support mechanisms. These components support predictive modeling, adaptive learning, and dynamic prediction in healthcare environments.

AI applications such as robotic-assisted procedures, disease prediction, and drug research have demonstrated significant potential in improving healthcare outcomes. A variety of studies have examined the advantages of AI technologies across multiple functional areas of smart healthcare. At the same time, the literature highlights challenges—particularly related to data quality, integration, privacy, and model reliability—that hinder full-scale adoption across healthcare systems.

The growing adoption of AI across industries has encouraged similar advances in healthcare. These technologies have the potential to transform many aspects of patient care and administrative operations. Multiple studies indicate that AI can match or even surpass human performance in tasks such as disease detection and risk assessment. However, despite promising results, AI still faces limitations in generalizability and

clinical applicability, necessitating further research to ensure safe and effective integration into clinical practice.

Healthcare systems worldwide require improvement to meet the increasing need for accessible, efficient, and high-quality patient care. Many individuals continue to face challenges in receiving adequate treatment for chronic illnesses, including heart disease, cancer, stroke, and diabetes. As healthcare moves toward value-based and patient-centered care, there is a growing need for approaches that provide personalized, cost-effective, and coordinated treatment. AI supports this goal by enabling physicians to access updated diagnostic information, detect abnormalities in imaging data, and compile comprehensive patient profiles based on clinical history and other relevant factors.

However, a significant portion of healthcare data nearly 80% exists in unstructured form. This makes it difficult for clinicians to access complete and organized information at the point of care. The lack of interoperability among systems, privacy concerns, and fragmented data storage add to the complexity of clinical decision-making. These issues highlight the need for modern data-driven approaches that facilitate precise, timely, and patient-focused care.

The increasing availability of healthcare data and advancements in big data analytics have enabled the development of successful AI applications for clinical decision support, diagnosis, and personalized treatment. Healthcare data is generated from numerous sources, including radiology, laboratory systems, wearable devices, sensors, physician notes, pathology reports, and clinical records. This diverse information contains important details such as demographics, medical history, family history, symptoms, test results, and treatment responses. Facilitating data sharing and integration across systems is essential for supporting proactive and preventive care.

Although access to integrated health data is often limited by privacy regulations and compatibility issues among medical devices, machine learning techniques can address many challenges once sufficient and authorized data becomes available. ML techniques have the capacity to identify patterns within data and support precision medicine by tailoring treatment decisions. Predictive models can, for example, assess the likelihood of hospital readmission in chronic disease patients and support timely interventions.

Discussions on artificial intelligence continue to expand across scientific and technical domains. Recent advances in image recognition, natural language processing, and speech analysis have raised interest in how these tools can enhance healthcare decision-making. Studies demonstrate that, in specific contexts, AI systems can achieve diagnostic performance comparable to experienced clinicians. However, realizing the full potential of AI in healthcare requires overcoming issues related to data quality, privacy, system compatibility, and responsible deployment.

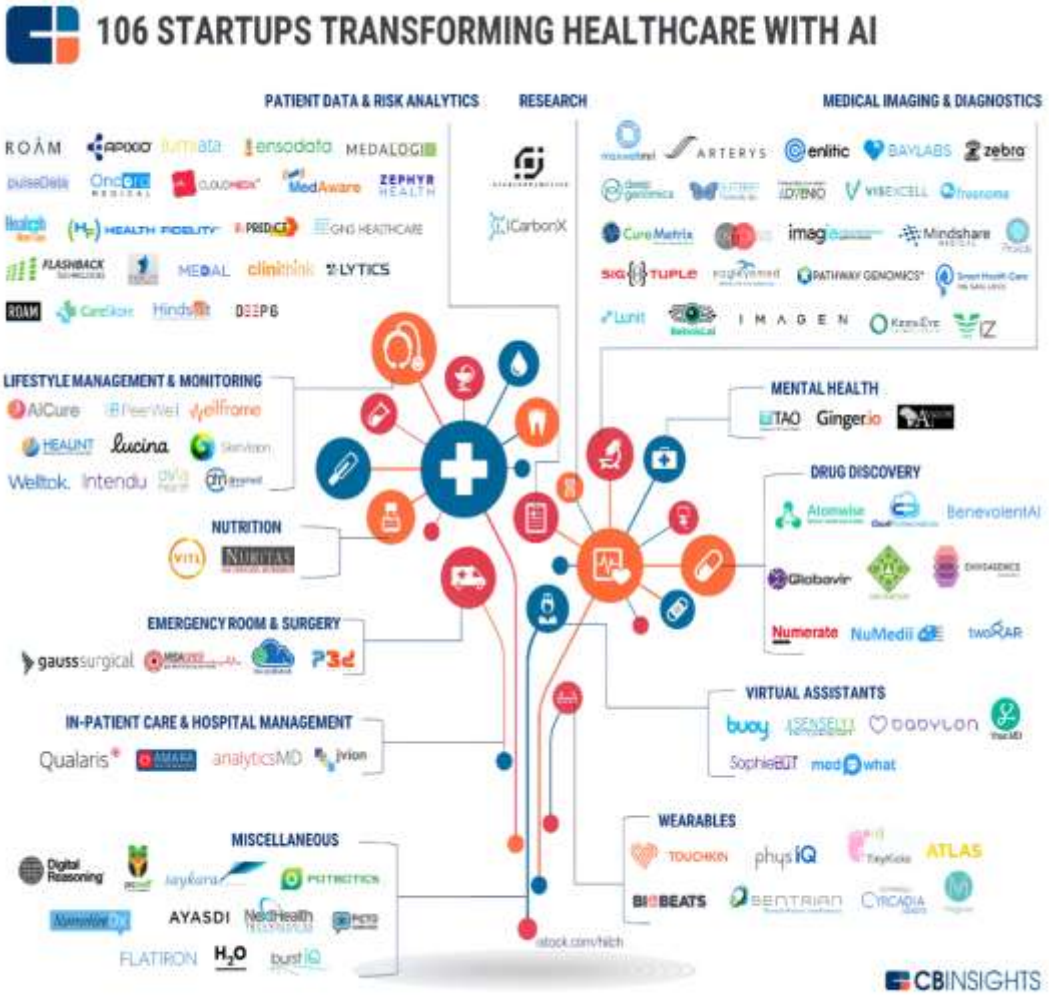

Figure 1.1: Robotics for New Medical Businesses. According to CB Insights (2016)

1.2 ARTIFICIAL INTELLIGENCE IN HEALTHCARE

Artificial Intelligence (AI) refers to computational systems capable of performing tasks that normally require human intelligence, such as learning from data, recognizing patterns, reasoning, and making informed decisions. Modern AI technologies—particularly machine learning, deep learning, and natural language processing—have become essential tools for analyzing the large and complex datasets generated in today's digital healthcare environment.

In healthcare settings, AI is used to interpret clinical information, support diagnostic decision-making, predict disease risk, and personalize treatment pathways. Its primary role is not to replace healthcare professionals but to augment their capabilities by providing accurate, data-driven insights that enhance clinical judgment and improve patient outcomes. By analyzing patterns in medical records, imaging data, laboratory results, and real-time physiological signals, AI systems can assist clinicians in identifying abnormalities earlier and more precisely than traditional methods.

AI applications in healthcare include early disease detection through medical imaging, outcome prediction using electronic health records, automated clinical documentation, and personalized treatment planning. These tools improve efficiency by reducing administrative workloads, enhancing diagnostic accuracy, and enabling faster decision-making. As the availability of digital health data continues to grow, AI is increasingly recognized as a critical component in achieving proactive, patient-centered, and value-based care.

Key Characteristics of AI in Healthcare

The implementation of AI-based smart healthcare solutions is defined by several core characteristics:

1. Big Data Processing and Analysis: Healthcare systems generate large volumes of structured and unstructured data from imaging equipment, wearable devices, laboratory systems, and electronic health records. AI models process this data to detect patterns, support clinical insights, and continuously learn from new information.

2. Augmented Intelligence:AI acts as a support system to enhance the capabilities of healthcare professionals. By providing accurate diagnostic suggestions, identifying risk factors, and managing vast patient datasets, AI assists clinicians in making timely and well-informed decisions. Full automation is limited to conditions where risks are minimal and human supervision is not critical.

3. Integration of Software and Hardware: AI in healthcare combines software technologies—such as machine learning, deep learning, and NLP—with medical hardware including imaging machines, monitoring devices, and robotic systems. Together, these components facilitate the analysis of both structured and unstructured data, supporting diagnosis, monitoring, and treatment planning.

1.3 NEED AND RATIONALE OF THE STUDY

Cardiovascular diseases (CVDs) remain the leading cause of morbidity and mortality worldwide, imposing substantial human, social, and economic burdens. Early and accurate prediction of heart disease risk plays a pivotal role in reducing adverse outcomes through timely clinical intervention, personalized treatment planning, and effective resource allocation. In many clinical settings, diagnosis and prognosis of cardiovascular conditions still rely on a combination of physician expertise, conventional risk scores, and limited diagnostic parameters. These approaches often underperform in two important ways: (a) they may not capture complex, nonlinear interactions among clinical variables, and (b) they are limited by incomplete, noisy, or heterogeneous data sources.

Rapid digitization of healthcare and the increasing availability of electronic health records (EHRs), physiological time-series (e.g., ECG), laboratory results, and imaging data create an opportunity to develop data-driven predictive models that augment clinical decision making. Machine learning (ML) and optimization-guided models can discover latent patterns and interactions not evident to traditional statistical tools. In the context of heart disease, such models can:

- Identify high-risk patients earlier than conventional screening tools, enabling preventive interventions and follow-up strategies.

- Improve prioritization and triage in resource-constrained settings (e.g., outpatient clinics, emergency departments), reducing delays to care.
- Personalize therapeutic recommendations by combining demographic, clinical, physiological, and laboratory indicators to produce individualized risk scores.
- Aid clinicians by offering consistent, reproducible risk estimates that complement clinical judgment, potentially reducing diagnostic variability and oversight.

However, effective translation of ML/AI solutions into clinical practice requires more than high accuracy on benchmark datasets: robustness to noisy / missing data, interpretable outputs, resistance to overfitting, and reliable generalization across populations are essential. This motivates the present study: the development of an integrated, optimized heart disease prediction framework that couples advanced feature selection and hybrid optimization algorithms (Modified Firefly Algorithm (MFA), Particle Swarm Optimization with Rough Sets (PSO-RS), Orthogonal Chicken Swarm Optimization (OCSO)) with strong classifiers (RBF-SVM, TSVM) to produce accurate, robust, and practically deployable prediction models. Using a real-world dataset of 303 patients**,** this research aims to deliver methods that are computationally efficient, interpretable for clinical use, and validated with rigorous performance metrics relevant to healthcare settings (accuracy, sensitivity, specificity, AUC, and clinical utility measures).

1.4 PROBLEM STATEMENT

Despite progress in ML for cardiovascular risk prediction, several persistent problems reduce the reliability and clinical readiness of existing models. These can be summarized as follows:

1. Feature-selection inefficiency and redundancy: Clinical datasets often contain many correlated and irrelevant variables (demographics, biochemical markers, comorbidities, ECG features). Traditional selection methods (filter/wrapper methods without global search) either retain noisy features or discard informative but subtle predictors, harming model performance.
2. Suboptimal hyperparameter tuning and model optimization: Many studies report models with good average performance but do not employ systematic

global optimization strategies for classifier kernels and parameters; this leads to models that are sensitive to initialization and dataset splits.

3. Poor handling of data uncertainty and incompleteness: Clinical records commonly contain missing values, measurement noise, and heterogenous scales. Existing approaches frequently use ad-hoc imputation and normalization methods that do not address uncertainty propagation through the learning pipeline.
4. Limited use of hybrid intelligent frameworks: Few studies combine advanced metaheuristic optimizers with robust classifiers in a modular pipeline that jointly optimizes feature selection and classifier parameters. Standalone ML algorithms or simple ensembles do not exploit the complementary strengths of metaheuristics (global search) and margin-based classifiers (SVM/TSVM).
5. Lack of interpretability and clinical explainability: High-performing black-box models are seldom accompanied by clinically interpretable explanations (feature importance, decision boundaries, local explanations), restricting clinical adoption.
6. Insufficient validation on diverse and adequately sized datasets: Research often relies on small or publicly available datasets with limited representativeness; cross-population generalizability and overfitting remain concerns.
7. Inefficient computational pipelines for real-time/near-real-time use: Methods that require heavy tuning, long training times, or large computation are impractical for deployment in hospital settings with constrained computing resources.

Therefore, the central problem addressed in this thesis can be formulated as:

How can we design a robust, accurate, and computationally efficient heart disease prediction system that (a) selects the most informative and stable features from noisy clinical data, (b) optimally tunes classifier parameters using advanced global search algorithms, and (c) provides interpretable, clinically actionable predictions validated on a real dataset of patients?

This research focuses on solving the above problem by integrating modified and hybrid metaheuristic optimizers (MFA, PSO-RS, OCSO) with RBF kernel SVM and

Transductive SVM classifiers, while emphasizing preprocessing, feature selection, and rigorous validation to ensure clinical relevance.

1.5 OBJECTIVES OF THE STUDY

Objective 1: To develop an efficient heart disease prediction system using modified firefly algorithm based radial basis function with support vector machine

To develop an efficient heart disease prediction model by integrating the Modified Firefly Algorithm (MFA) for optimal feature selection and parameter tuning with an RBF-SVM classifier. The objective is to enhance diagnostic accuracy and clinical reliability. Model performance will be assessed using accuracy, sensitivity, specificity, AUC, and other standard evaluation metrics.

Objective 2: To configuration upgraded expectation of coronary illness utilizing molecule swarm advancement and harsh sets with transductive help vector machines grouped

To design a heart disease prediction framework that combines Particle Swarm Optimization with Rough Set theory (PSO-RS) for feature reduction and applies Transductive SVM (TSVM) for classification. The objective is to improve uncertainty handling, utilize semi-supervised learning capability, and achieve more robust generalization on the available dataset.

Objective 3: To develop Proficient System to Identify Heart Diseases with the Aid of Artificial Intelligence and Soft Computing Techniques

To investigate OCSO as an advanced metaheuristic technique for optimizing classifier parameters. This objective involves comparing the performance of OCSO with MFA and PSO-RS in terms of convergence behavior, robustness, and predictive accuracy under identical experimental conditions.

Objective 4: To Assist Physicians in Predicting and Diagnosing Cardiovascular Diseases at An Early Stage Effectively and Accurately

To develop a complete heart disease prediction pipeline comprising data preprocessing, normalization, feature selection, optimization, and classification. This includes evaluating model interpretability, reliability, and clinical applicability to ensure that the proposed system can effectively support physicians in early diagnosis and clinical decision-making.

1.6 APPLICATIONS OF AI IN HEALTHCARE

Artificial Intelligence has emerged as a transformative technology in modern healthcare by enhancing diagnostic accuracy, improving the efficiency of clinical workflows, supporting therapeutic decision-making, and enabling personalized treatment strategies. Its applications span multiple domains of medical science, bringing significant improvements in disease prediction, patient monitoring, drug development, and clinical operations.

AI for Drug Discovery

AI significantly accelerates the drug discovery process by automating target identification, predicting molecular interactions, and identifying potential compounds for clinical investigation. Machine learning models analyze vast biochemical datasets to repurpose existing drugs and explore novel therapeutic candidates. This reduces development timelines and costs compared to conventional laboratory-driven processes. Several leading pharmaceutical companies have adopted AI-driven platforms for oncology, metabolic disorders, and immunotherapy research, demonstrating its expanding role in next-generation drug development.

AI in Clinical Trials

Clinical trials involve extensive data management, patient selection, and monitoring, often making them time-consuming and costly. AI supports these activities by automating data processing, identifying suitable patient cohorts, predicting trial outcomes, and improving trial design through real-world data (RWD) analysis. Intelligent systems can clean, aggregate, and code clinical data with greater accuracy, reducing manual errors. AI-driven models also enhance cooperation among research

institutions by enabling secure data sharing, model transfer, and cross-institutional analytics—thereby accelerating medical research.

AI for Patient Care

AI enhances patient care by analyzing clinical histories, identifying high-risk individuals, and supporting personalized treatment pathways. Clinical intelligence systems evaluate electronic health records to provide actionable insights for clinicians. Examples include early detection tools for maternal health risks, AI-assisted monitoring systems for chronic disease management, and predictive models that alert clinicians to potential complications. These systems strengthen early intervention and improve overall quality of care.

Healthcare Robotics

AI-driven robotic systems assist in rehabilitation, surgery, and patient support. Exoskeleton robots improve mobility for individuals with spinal or neurological impairments, while smart prosthetics offer enhanced precision and functional capability through sensor-driven control. Robots equipped with AI also support post-surgical recovery, physiotherapy, and assistive tasks, thereby improving patient autonomy and reducing caregiver burden.

Genomic and Data-Driven Medicine

Genomic data analysis supported by AI enables personalized medicine by uncovering genetic markers associated with disease susceptibility and treatment response. Wearable health devices and biosensors continuously collect physiological data, allowing AI algorithms to anticipate medical conditions, predict disease progression, and recommend lifestyle or therapeutic adjustments. These advancements help clinicians offer more targeted and individualized care.

AI-Enabled Diagnostic Tools

AI-powered diagnostic devices such as digital stethoscopes, smart imaging tools, and pattern-recognition systems analyze physiological signals and medical images with high accuracy. For example, AI can detect subtle abnormalities in cardiovascular

sounds, radiology images, or laboratory data that may not be easily visible to human experts. These tools enhance diagnostic precision, particularly in remote or underserved regions where specialist availability is limited.

1.7 TYPES OF ARTIFICIAL INTELLIGENCE USED IN HEALTHCARE

Artificial Intelligence in healthcare is implemented through several computational approaches that support data analysis, diagnosis, prediction, and clinical decision-making. The most widely used AI types in medical applications include the following:

1. Natural Language Processing (NLP)

NLP enables computers to understand and process clinical text such as electronic health records, physician notes, laboratory reports, and discharge summaries. It assists in automated documentation, information extraction, clinical coding, and decision-support tasks, improving workflow efficiency and reducing manual effort.

2. Machine Learning (ML)

Machine learning algorithms learn patterns from structured clinical data, enabling disease prediction, risk assessment, and treatment optimization. ML is widely applied in heart disease prediction, medical imaging, and drug safety analysis. Techniques such as decision trees, SVM, and ensemble methods form the backbone of ML-based healthcare analytics.

3. Deep Learning (DL)

Deep learning utilizes multi-layer artificial neural networks to analyze high-dimensional data such as medical images, physiological signals, and genomic sequences. Convolutional neural networks (CNNs) support diagnostic imaging, while recurrent neural networks (RNNs) and LSTM models assist in ECG signal interpretation and temporal health data analysis.

4. Rule-Based Expert Systems

These systems use predefined medical rules (“if–then” conditions) to support clinical decision-making. They provide interpretable recommendations for diagnosis and

treatment pathways, especially in settings where historical clinical knowledge is crucial.

5. AI-Driven Automation

Robotic Process Automation (RPA) and intelligent automation streamline administrative tasks such as appointment scheduling, billing, insurance verification, and report generation. This reduces workload on healthcare staff and enhances operational efficiency.

1.8 SCOPE OF THE STUDY

The scope of the present study is clearly defined to ensure focused investigation, avoid unnecessary expansion, and establish the specific boundaries within which the proposed heart disease prediction framework is developed and evaluated. This research concentrates on the design, optimization, and assessment of machine learning and soft computing–based prediction models using the available patient dataset and selected computational methodologies. The scope outlines the components included for analysis as well as the areas intentionally excluded from this thesis.

1.8.1 Included Scope

Dataset and Study Population

This study uses the available dataset consisting of 303 patient records related to heart disease. Detailed attribute descriptions, distributional characteristics, missing value patterns, and ethical considerations regarding anonymization and data usage will be presented in Chapter 3. All preprocessing steps applied to this dataset fall within the defined scope of the research.

Data Preprocessing Procedures

The study includes data cleaning, handling of outliers, missing value imputation through statistical or model-based techniques, normalization or standardization, and encoding of categorical attributes. These steps ensure that the dataset is suitable for the development of reliable prediction models.

Feature Selection and Attribute Reduction

The research covers the application of multiple feature selection and reduction techniques, including the Modified Firefly Algorithm (MFA), Particle Swarm Optimization with Rough Sets (PSO-RS), and baseline attribute selection measures such as mutual information and recursive feature elimination. The objective is to identify compact and clinically meaningful subsets of predictive features.

Optimization Algorithms

This study investigates MFA, PSO-RS, and Orthogonal Chicken Swarm Optimization (OCSO) for optimizing both feature subsets and classifier parameters. Their performance, computational efficiency, and stability will be comparatively analyzed.

Classification Models

The primary classification models included in the study are RBF-SVM (optimized using MFA and OCSO) and TSVM (optimized using PSO-RS). For benchmarking purposes, classical machine learning methods such as logistic regression, standard SVM, and random forests may also be employed to provide comparative evaluation.

Evaluation Metrics and Validation Strategy

The performance of the proposed models will be assessed using stratified k-fold cross-validation, holdout validation, and repeated experiments where required. Standard evaluation indicators such as accuracy, sensitivity, specificity, precision, recall, F1-score, AUC, and calibration measures will be used, along with statistical significance tests like paired t-tests or Wilcoxon signed-rank tests.

Interpretability and Clinical Reporting

The study includes interpretability analysis such as feature ranking, rule-based or surrogate interpretive models, and discussion of clinical relevance of selected attributes. Recommendations for presenting model outputs to healthcare professionals are also part of the scope.

Implementation and Feasibility Considerations

The study discusses computational requirements, algorithmic complexity, runtime aspects, and potential integration pathways into clinical decision-support environments. These considerations help evaluate the practical applicability of the developed system.

1.8.2 Excluded Scope

Large-scale Clinical Trials or Real-time Deployment

The study is limited to retrospective analysis using the provided dataset. Prospective clinical trials, hospital-level deployment, or real-time integration into medical systems are outside the scope of this thesis.

High-Dimensional Genomic or Raw Imaging Data

Unless such data are already represented within the available 303-sample dataset, high-dimensional genomic information or raw medical imaging (such as DICOM image processing) is excluded. Only structured or derived features present in the dataset will be utilized.

Regulatory Certification and Legal Compliance Processes

While ethical data usage principles are acknowledged, detailed regulatory approval activities required for medical device certification lie beyond the present study's scope.

Commercial Product Development

The research focuses on algorithmic development and experimental validation rather than commercialization or full production deployment of a medical device.

1.8.3 Assumptions and Limitations

This study assumes that the 303-patient dataset is representative of the population for which the model is intended. As such, generalization to other populations should be approached cautiously and validated separately. The moderate sample size may limit the statistical strength for rare subgroups. Additionally, hybrid optimization models

may achieve high predictive accuracy but may require supplementary interpretability measures to ensure clinical acceptance.

1.9 THE FUTURE OF AI IN HEALTHCARE

Dr. Jehi states that research remains the area where artificial intelligence in healthcare has the greatest potential. Through her clinical experience, she observed that AI still has much to teach the medical community by uncovering patterns that humans cannot easily perceive. As an expert studying epilepsy surgery, Dr. Jehi highlights how machine learning is transforming traditional clinical decision-making.

Previously, surgeons relied on multiple clinical tests—brainwave recordings, neuroimaging, and specialist interpretations—to determine the brain region responsible for seizures. These decisions were largely based on individual clinical experience, which limited the ability to generalize or compare across large patient groups. As a result, treatment choices for new patients were made with limited collective knowledge.

Machine learning has now enabled the aggregation of patient data into unified analytical systems. By centralizing and analyzing large volumes of clinical information, AI helps physicians better understand disease patterns, compare treatment outcomes, and make more informed decisions. Importantly, the clinical tests themselves have not changed; rather, AI enhances the depth of insights extracted from these tests.

The goal of ongoing research is to develop more accurate AI-based prediction models to support medical and surgical decisions in epilepsy and other conditions. Researchers are working on simplifying these models so they can be integrated into routine clinical workflows. Using machine learning, Dr. Jehi and her team have identified indicators associated with surgical complications or recurrence, and automated systems for detecting and localizing abnormal brain tissue are also in development.

Another area of interest is understanding how a patient’s genetic makeup and brain characteristics influence seizure behavior and long-term surgical outcomes. Emerging evidence suggests that genetics plays a significant role in determining the success of epilepsy interventions. With AI and ML, researchers aim to study larger patient cohorts

to uncover deeper relationships between clinical features, genetics, and treatment response.

Overall, the future of AI in healthcare lies in combining clinical expertise with large-scale data analysis to deliver more precise, personalized, and consistent care. Continued research and model refinement will help translate these innovations into practical tools that can meaningfully support clinical decision-making.

1.10 CHALLENGES AND OPPORTUNITIES OF USING AI FOR IMPROVING HEALTHCARE

Artificial Intelligence (AI) has become an essential component of modern healthcare, offering improved diagnostic accuracy, faster decision support, and enhanced efficiency. However, despite these advantages, several challenges must be addressed for AI systems to be integrated effectively and responsibly within healthcare settings. These challenges arise from ethical, technical, organizational, and clinical factors and directly influence the reliability, acceptance, and long-term sustainability of AI-driven prediction systems.

A major concern relates to ethical and privacy issues, particularly around patient autonomy, informed consent, and secure handling of sensitive medical information. Since AI systems rely on large, high-quality datasets, ensuring privacy-preserving data sharing and compliance with regulatory frameworks remains a critical requirement. Healthcare data is often fragmented, incomplete, and non-standardized, leading to difficulties in achieving high model performance, interoperability, and reliable generalization across patient populations.

The rapid expansion of biomedical knowledge further challenges clinicians, as the pace of new scientific insights exceeds their ability to manually interpret and apply them in practice. AI tools can support this process, but their successful integration requires robust validation and user-friendly interpretability. Additionally, the rise in multimorbidity complicates the clinical decision-making process, as traditional single-disease guidelines are insufficient, creating a need for advanced AI models capable of managing complex interactions.

Another major issue concerns the fairness, transparency, and explainability of AI models. Bias in training datasets may produce unequal or inaccurate predictions for certain demographic or clinical groups. Clinicians also require clear and interpretable outputs to maintain trust and ensure appropriate use of model recommendations. Responsibility and accountability for AI-generated decisions remain ambiguous, raising concerns about professional liability, especially when AI errors might affect patient outcomes.

On the technical side, many healthcare institutions lack the required computational infrastructure and interoperable electronic health records needed for seamless AI deployment. The “black-box” nature of advanced machine learning models further limits adoption, as clinicians and administrators prefer systems that provide traceable reasoning. Moreover, limited digital literacy within healthcare environments affects the willingness and ability of practitioners to adopt AI-based tools.

Despite these challenges, AI presents substantial opportunities for transforming healthcare. Predictive analytics can support early detection, risk stratification, and timely intervention, thereby improving patient outcomes and reducing healthcare costs. AI-driven decision support systems can enhance the efficiency of clinical workflows, enable personalized treatment recommendations, and assist in managing complex or high-volume data. When properly validated and implemented with strong ethical safeguards, AI has the potential to support clinicians, strengthen diagnostic accuracy, and contribute to more reliable and patient-centered healthcare systems.

Fig. 1.2 Challenges in healthcare AI described.

1.11 THE BENEFITS OF USING AI IN HEALTHCARE AND HOSPITALS

Artificial intelligence (AI) refers to the capability of computational systems to perform tasks that traditionally required human intelligence, such as speech recognition, decision-making, and language translation. Machine Learning (ML), a major subfield of AI, enhances this capability by enabling systems to learn from large datasets and solve complex problems using data-driven algorithms. Together, AI and ML significantly improve the efficiency and effectiveness of healthcare processes by enabling rapid data processing, pattern recognition, and evidence-based insights.

AI technologies now support various areas of medical practice such as diagnostic imaging, neurology, emergency care, and administrative services. These systems analyze large clinical datasets in the background and enhance patient care even before

individuals arrive at a healthcare facility. Their integration encourages clinical teams to reconsider existing workflows and adopt more efficient, data-driven approaches.

In healthcare, AI systems have demonstrated the ability to interpret certain forms of medical imaging—such as CT and MRI scans—with high accuracy. AI tools, including emerging generative models, continue to evolve and show potential for contributing to clinical decision-support systems, diagnostic interpretation, and predictive modelling.

AI is also being applied to accelerate biomedical research by enabling faster analysis of high-dimensional data such as genetic sequences, molecular interactions, and physiological markers. The increasing availability of biological data supports advanced AI-driven methods that improve understanding of disease mechanisms and inform diagnosis, treatment planning, and follow-up care. As healthcare continues to evolve, these technological developments require clinicians to adopt new skills related to data interpretation and computational tools.

AI-driven methods enhance diagnostic precision by offering rapid and accurate analysis of medical images, allowing early identification of conditions such as fractures, cancer, and vascular abnormalities. In time-sensitive cases such as stroke, accelerated evaluation supports faster clinical decision-making and improved outcomes.

Rapid Diagnosis

AI algorithms can process imaging data with high speed and accuracy, enabling early detection of abnormalities and reducing diagnostic delays. This leads to quicker initiation of treatment, reduced patient anxiety, and improved overall satisfaction with care.

Assistance in Surgery

AI-enabled robotic systems support minimally invasive surgical procedures by enhancing precision and offering real-time feedback based on intraoperative data. These systems assist clinicians in navigating complex anatomical structures and reducing surgical risks.

Improved Accessibility of Care

AI-enabled telemedicine platforms expand access to healthcare services for individuals in remote or underserved areas. Through virtual consultations, automated symptom guidance, and remote monitoring, patients can receive timely clinical support without geographical limitations.

Patient Support and Self-Management

AI-powered chatbots and virtual assistants help patients manage health-related tasks by providing reminders, answering queries, and suggesting self-care practices. These tools support patient engagement and continuity of care outside traditional clinical settings.

1.12 USE OF AI IN HEALTHCARE

What we have discussed so far is how AI is bringing about change and improvement in the healthcare sector. Next, we'll look at several practical applications of AI in medicine:

Collaborative effort and making choices: It is essential for healthcare practitioners to work together in teams in today's healthcare systems. This calls for open and honest communication, team decision-making, coordinated activities, and regular evaluations of success. As mentioned in, AI chatbots may help with medical appointment scheduling and coordination, reminders, and symptom-based condition notification to clinicians.

As previously said, technological breakthroughs such as health monitoring systems powered by artificial intelligence may greatly benefit the elderly. These systems guarantee prompt delivery of treatment, free up healthcare practitioners to give more comprehensive care outside of regular office hours, and encourage self-management. For instance, as mentioned in, sensor technology may streamline self-monitoring for heart failure patients by using user-friendly gear. There are a number of health-related technologies that might manage medical students' and practitioners' laboratory procedures. As stated in, a virtual reality simulator may help inexperienced surgeons hone their skills in a controlled environment, where they can rehearse treatments in advance and plan for any contingency, ultimately leading to safer, more accurate

surgeries. Furthermore, as previously said, the public's view of healthcare is evolving due to technological developments. As previously mentioned, AI is also finding applications in the control of processes, image analysis, virtual assistants, robotic surgery, and clinical decision support. Possibilities in artificial intelligence are detailed in

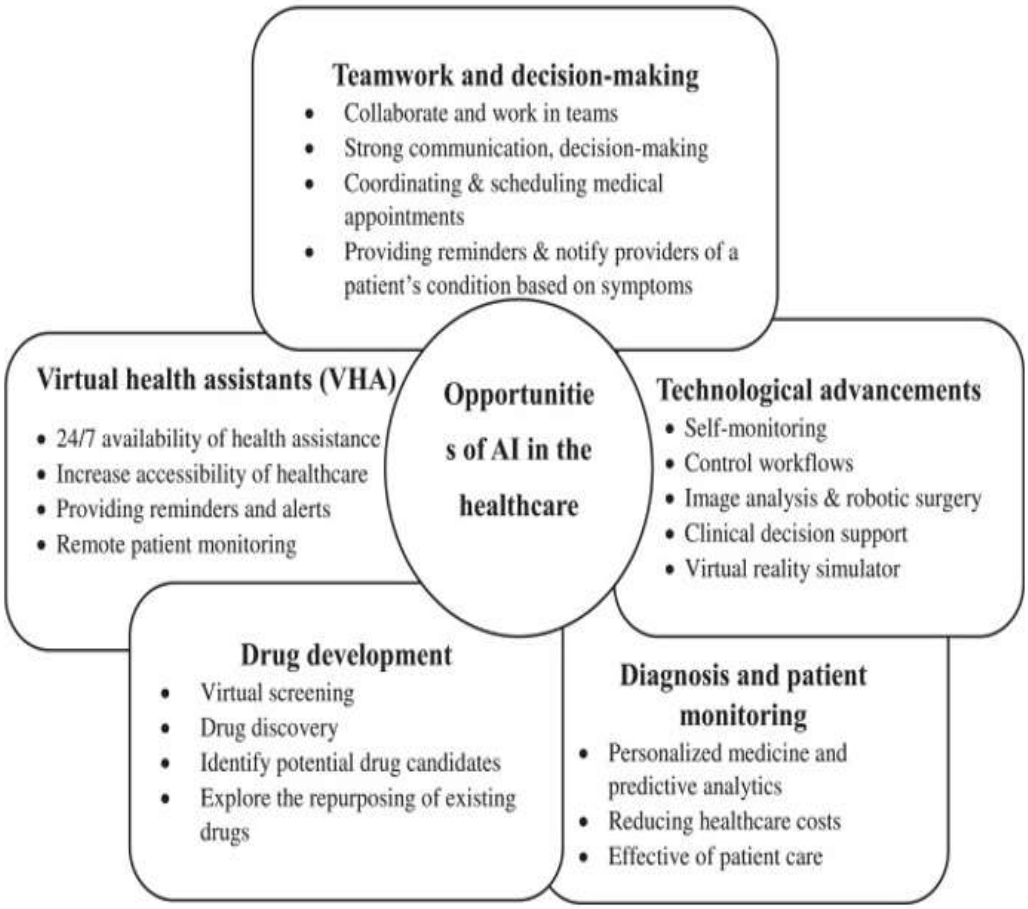

Fig. 1.3 Potential applications of artificial intelligence in healthcare.

Better patient outcomes and lower healthcare costs are the benefits of early illness identification and progression tracking made possible by artificial intelligence (AI) in the context of personalized medicine and predictive analytics. Also, AI may help with therapy development, which is great for healthcare innovation and research. There are several unique prospects for the use of AI to greatly enhance the efficacy and efficiency of patient care. To better understand which patients are most likely to experience a decline in health and which ones are more at risk of problems, predictive analytics may

be used. Healthcare practitioners may enhance patient outcomes and forestall the development of serious conditions by acting early.

Pharmaceutical research and development: AI offers enormous promise for streamlining and improving pharmaceutical research and development. The possibility of using virtual screening to examine massive volumes of data on medication interactions and discover novel therapeutic targets is one such option. This has the potential to significantly shorten the time it takes to find new drug candidates, which in turn may reduce the cost of drug development. Also, AI may look at data on medication prospects to find the best compounds to develop further and look into ways to repurpose current pharmaceuticals for new applications.

1. AI in Drug Discovery and Development

There would be no healthcare system without the pharmaceutical business. Being a leading beneficiary, their work in medicine development enables clinicians to treat patients, which has the potential to save lives.

One of the most prominent applications of artificial intelligence (AI) in healthcare is the pharmaceutical industry's heavy investment in R&D for the purpose of finding and creating new medications. Artificial intelligence (AI) technology may greatly improve the speed and efficiency of the pharmaceutical drug development pipeline, which is now characterized by a high reliance on human labor.

The first medication to be completely AI-designed has made it to human clinical trials, thanks to generative AI-driven biotech businesses like Insilco Medicine.

Here are a few important points that point to the significant role that AI may play in the pharmaceutical industry:

- Artificial intelligence systems can sift through mountains of biological data, such as genetic sequences, molecular models, and results from clinical trials, to find new medicines.

- It is possible to precisely identify the disease targets using AI-based predictive modeling tools. It makes it easier to analyze patient data, such as genomes and clinical records, in order to create customized medicines that are specific to each patient.

- By allowing the repurposing of pharmaceuticals via the examination of current drugs' chemical structures and biological effects, this method significantly reduces the time spent producing new treatments.

- AI models may analyze the chemical qualities and biological interactions of drug candidates, anticipate their possible side effects, and eliminate the possibility of safety concerns during clinical trials, all with the goal of lowering the risk of toxicity and bad consequences.

- Finally, artificial intelligence is changing the face of clinical trials by making "digital twins," or digital copies, of each patient. All of them act as living computer models that mimic normal and abnormal human physiological and pathological functions. It allows for predictive and tailored insights all the way through a clinical study.

2. AI in Personalized Medicine

Precision medicine, sometimes called customized medicine, is an alternative to traditional medicine's one-size-fits-all approach. It entails creating a unique treatment plan for each patient by examining their medical history, genetic information, lifestyle choices, and other relevant data.

Unlike conventional medicine, which primarily aims to alleviate symptoms, precision medicine tailors its approach to each patient based on their unique requirements. To enhance treatment results while minimizing unwanted effects, a data-driven strategy is used, which involves assessing numerous factors concurrently.

In addition, by using wearables or remote sensors, real-time monitoring provides ongoing insights on the patient's health condition and response to therapy. Improving patient care is made possible by prompt interventions and modifications to treatment regimens.

3. AI in Medical Imaging

When evaluating a wide range of medical images—including X-rays, MRIs, CT scans, ultrasounds, and more—the precision of the diagnosis is crucial. Even while radiologists are quite good at interpreting these pictures, they are still human and may

make mistakes. Furthermore, this procedure may be laborious, particularly in cases when picture anomalies need meticulous examination to precisely pinpoint the root cause.

To teach AI to correctly identify patterns or irregularities, however, hundreds of photos with varied issues are sent into the system. They won't be able to take the position of radiologists entirely, but they may help them save time and be more precise in their diagnosis by catching every information that has to be considered.

4. AI in Genomic Medicine

With genetic medicine, AI elevates individualized therapies to a whole new level. To begin, genetic data is very complicated and contains a large quantity of information; thus, sophisticated computing techniques are required for its analysis.

Consumption of time will persist even with these instruments. AI-powered algorithms sift through the available genomic data in search of genetic markers linked to certain features, illnesses, or treatment reactions.

In addition, clinicians may use genetic data to create prediction models that AI can use to determine an individual's susceptibility to particular illnesses or the effectiveness of certain therapies. This enables them to suggest better lives or targeted medical treatments to lower the likelihood of certain illnesses.

5. AI in Robotic Surgery

My imagination immediately goes to a sci-fi scenario where surgeons use robotic arms controlled by a computer to do less intrusive and more accurate procedures on patients. I am no longer dreaming about this. While operating from a control console, surgeons are assisted by surgical robots such as the da Vinci Surgical System.

Both patients and physicians favor minimally invasive techniques that are precise because they increase the success rate of surgery. Improved patient outcomes are the end result of robotic surgical systems driven by artificial intelligence that allow surgeons to execute complicated operations with more accuracy, efficiency, and safety.

6. AI in Patients' Assistance

Virtual assistants like this simplify healthcare for everyone involved. The function of a virtual assistant is as follows:

- They aid patients by responding to inquiries, serving as gentle reminders, and giving emotional support.
- They help with appointment scheduling, patient record organization, and enabling easy access to medical information, among other things.
- They make healthcare information and resources easily accessible to patients.
- Using the data collected from your wristwatch, they may assess your activity levels and provide advice on how to maintain a healthy lifestyle.

7. AI in Oncology

In addition to other uncommon illnesses, cancer is one that might benefit from AI's use in both diagnosis and medication development. An example of an AI application in oncology might be:

- Recognition (of kinds, stages, and health issues) or accurate/early diagnosis using precise analysis of medical pictures (e.g., CT scans, X-rays, MRIs, and more).
- Creating an individualized treatment strategy by sifting through mountains of data on a patient's health, genetics, pathology findings, and more.
- Analytics that may foretell the patient's reaction to drugs, side effects, and other factors related to the chosen chemotherapy treatment or alternatives.
- Tailored suggestions for cancer treatments, including kinds of treatments, dose quantities, and more, to maximize therapy efficacy with minimum risk of adverse effects and drug overexposure.

8. AI in Remote Patient Monitoring (RPM)

Using AI for RPM is like knowing what's happening in your hour, except instead of you, healthcare providers can know your vitals like blood pressure, respiratory rate,

and more from anywhere thanks to the internet of things. We've already seen the power of the internet of things.

A smartwatch or other wearable device may monitor the vitals that are specifically being monitored by a healthcare professional, as well as general vitals that provide a picture of the patient's health as a whole, such as blood pressure and heart rate. An AI-powered mobile app is synced with the devices and scours the gathered data for any suspicious patterns.

The tool also provides doctors with access to these datasets. The app will notify the doctor if it detects anything suspicious, either via pattern analysis or any unexpected increases from the patient's specific baselines. This will allow the doctor to promptly address the matter.

9. AI in Mental Health Support

When it comes to AI and healthcare, it's not just about physical health; AI has also proven very beneficial for mental health.

Many people with mental health concerns go undiagnosed or untreated until it's too late, which may have devastating consequences, including terrible results like suicide. People with mental health issues may suffer in silence for a variety of reasons, including a lack of understanding about the gravity of their disease and the stigma associated with seeking treatment.

Consequently, family ones and healthcare professionals may be ignorant of the individual's challenges until the problems reach a critical stage, at which time it is too late to do anything about it.

A few key points illustrating the applications of AI in mental health are as follows:

- Algorithms trained by machines may spot trends that can indicate mental health issues like bipolar disorder, depression, or anxiety, allowing for quicker treatment.

- People dealing with mental health issues have constant access to virtual assistants and chatbots driven by artificial intelligence.

- Artificial intelligence systems assess a person's propensity for suicidal thoughts and actions by analyzing risk variables such social isolation, drug misuse, past suicide attempts, and behavioral changes.

- Natural language processing tools powered by artificial intelligence examine audio and text recorded during counseling, support group, and therapy sessions to derive valuable insights on patients' attitudes, feelings, and treatment outcomes.

10. AI in Clinical Documentation

In a hectic medical practice, doctors waste time that might be better spent diagnosing patients by typing or, worse, writing down details about their symptoms, medical history, and possible treatments.

The doctor-patient communication may be transcribed and analyzed by online or mobile apps or even search engine extensions that use artificial intelligence to suggest possible treatment plans.

Our team recently developed on a feature called "Scribe" that incorporates AI into clinical documentation as part of our project Sully.ai, an AI-powered all-in-one tool for physicians. It goes so far as to provide (or rather produce) a clinical strategy for the physicians after the diagnosis is made.

11. AI in Fraud Detection

With its share of false invoices, needless treatments, and other forms of insurance claim fraud, the healthcare industry is a major player on a worldwide scale. Medical providers and hospitals submit hundreds of claims to healthcare insurance companies daily for services rendered to patients. These firms are finding it more challenging to identify warning signs due to the high volume.

In order to identify any fraud, AI-powered fraud detection software compares the provided facts to the claim and looks for warning signs. When it detects claims that don't add up, it notifies the relevant insurance agency to look into them further. The insurance firm saves a significant amount of money by taking this proactive strategy.

1.13 CHAPTER SUMMARY

Chapter 1 provides the foundational context for the study by outlining the growing role of Artificial Intelligence (AI) in modern healthcare and its potential to transform diagnostic processes, predictive modelling, and patient care. The chapter highlights the increasing burden of cardiovascular diseases and the limitations of traditional diagnostic approaches, establishing the need for more accurate, data-driven prediction systems. It also emphasizes why AI, with its ability to identify complex nonlinear relationships in clinical data, is well-suited for improving heart-disease prediction accuracy and supporting early clinical interventions.

The chapter clearly identifies the rationale of the study, the problem addressed, and the specific research gaps in existing literature—such as inefficient feature selection, incomplete handling of uncertain data, limited parameter optimization, and lack of interpretable models. Based on these gaps, the objectives of the research are defined, focusing on developing optimized predictive models using hybrid algorithms like MFA, PSO-RS, and OCSO integrated with RBF-SVM and TSVM classifiers.

Furthermore, the chapter discusses the scope, assumptions, limitations, challenges, and benefits associated with AI applications in healthcare. It also introduces the wide range of AI-enabled tools and techniques being used across clinical domains. Overall, Chapter 1 establishes the motivation, significance, and direction of the thesis, providing a clear platform for the detailed methodology and experimental design presented in subsequent chapters.

CHAPTER-2

LITERATURE REVIEW

This chapter presents a concise and structured review of existing research on artificial intelligence (AI) and machine learning (ML) in healthcare, with emphasis on predictive modelling for cardiovascular disease. The selected studies demonstrate how AI has evolved from basic computational techniques to advanced clinical decision-support systems capable of analyzing complex medical data. By examining past work on diagnostic applications, predictive analytics, optimization algorithms and healthcare informatics, this review identifies the strengths, limitations and research gaps in current AI-based healthcare solutions. These insights form the foundation for the proposed optimized predictive models developed in this study.

Jiang et al. (2017) describe artificial intelligence (AI) as the emulation of human cognitive capabilities such as learning, reasoning, and decision-making through computational models. Their work documents some of the earliest AI systems deployed in oncology, neurology, cardiology, and stroke care, demonstrating that AI can analyse structured clinical databases as well as unstructured information such as radiology reports and physician notes. The authors show that these systems significantly improve diagnostic confidence and prognostic estimation while reducing manual workload. However, they also emphasise key challenges including model interpretability, integration into clinical workflows, and the need for continuous validation to ensure reliability across diverse patient populations.

Yu et al. (2018) provide a broad and foundational review of advances in AI and their biomedical applications, noting how progress in machine learning algorithms, digital records, and high-resolution biomedical sensors has enhanced clinical decision support. Their work highlights that AI can uncover subtle, nonlinear relationships in high-dimensional datasets—an essential capability for complex diseases such as cardiovascular disorders. Nevertheless, the authors caution that large-scale adoption requires addressing regulatory constraints, ethical considerations, data privacy issues, and the financial burden associated with deploying AI technologies in clinical environments.

Murali et al. (2018) position AI as a rapidly expanding subfield of computer science that increasingly outperforms human experts in specialised diagnostic tasks. Their review underscores applications in neurological disorders, diabetes, cardiovascular disease, and various cancers, showing how AI models can detect nuanced trends within patient data and identify early indicators of disease. However, they argue that despite the promising performance, robust clinical validation and transparent model behaviour are critical prerequisites for safe clinical integration.

Haleem et al. (2019) identify five major AI technologies—machine learning, natural language processing, robotics, expert systems, and deep learning—and summarise their ten key applications within healthcare. These include clinical decision support, personalised therapy selection, infection surveillance, hospital workflow optimisation, and predictive analytics. Their findings show that AI enhances decision-making in complex clinical situations, but they also emphasise the need for clinician acceptance, training, and strong governance frameworks to ensure responsible deployment.

Bohr et al. (2020) highlight how big data and machine learning permeate modern healthcare, supporting tasks across the full value chain—from patient registration and administrative documentation to image analysis, predictive diagnostics, and ambient assisted living. They argue that AI systems augment human capabilities rather than replace clinicians, enabling them to make faster, more informed decisions while reducing cognitive overload.

Tadiboina et al. (2021) extend this view by examining AI adoption across the life sciences industry, healthcare providers, and insurance payers. They describe the diverse uses of AI in administrative automation, patient engagement, therapeutic adherence monitoring, diagnostic recommendations, and claims management. Their work concludes that while AI will primarily complement healthcare professionals, several operational roles will undergo substantial transformation due to automation and predictive analytics.

Reddy et al. (2020) describe the broader ecosystem of AI in healthcare, discussing its applications in medical diagnosis, population health monitoring, genomic prediction, and administrative optimisation. They report significant investments by governments, universities, and technology firms into AI-driven health innovations. However, they

also emphasise that many stakeholders still lack clarity regarding AI's limitations, operational requirements, and ethical implications, which remain barriers to widespread adoption.

G.M. et al. (2021) define healthcare AI as "augmented intelligence," highlighting its purpose of supporting rather than replacing clinicians. Their review covers AI applications for diagnosis, prognosis, and therapy planning, as well as advanced algorithms for medical image processing, feature extraction, and patient-care optimisation. Their findings support the growing consensus that AI can enhance diagnostic speed, reduce variability in interpretation, and improve patient outcomes when appropriately integrated.

Raj et al. (2023) present AI as a rapidly maturing discipline capable of transforming multiple domains of healthcare, including cancer detection, neurological assessment, cardiovascular disease prediction, and diabetes management. They also demonstrate that AI can accelerate drug discovery pipelines, clinical trials, and personalised treatment recommendations by efficiently analysing large biomedical datasets.

Within this broad landscape, the present research specifically targets AI-based predictive systems for cardiovascular disease, focusing on optimised SVM and TSVM classifiers enhanced through Modified Firefly Algorithm (MFA) and Particle Swarm Optimization (PSO) for attribute reduction. Unlike many of the reviewed studies that discuss AI generally, this work develops and evaluates hybrid optimisation-driven classification models using structured clinical datasets (1000–5000 records from the Cleveland Heart Disease Dataset). The goal is to enhance diagnostic accuracy, minimise false positive/negative rates, and provide clinicians with a reliable, interpretable tool for early heart disease detection and personalised risk assessment.

Jimma et al. (2023) conducted a bibliometric analysis of 5,019 AI-in-healthcare publications from 2000 to 2021 and reported an exponential surge in research output after 2012. This rapid growth was driven by advances in machine learning, electronic health records, natural language processing and the increasing availability of clinical data. They note that major disease areas—COVID-19, diabetes, mental health and cancer—dominate global publications, demonstrating AI's expanding relevance across

clinical domains. Their findings show how AI research has transitioned from isolated pilot experiments to a core component of mainstream medical innovation.

Amit et al. (2022) reviewed more than 4,000 AI-healthcare papers published in 2021, mapping the field's evolution into three major hotspots: predictive analytics, medical imaging and clinical decision support systems. Their analysis indicates that AI models increasingly focus on early detection, risk stratification and automated diagnosis—all of which align closely with the predictive modelling goals of the present research. The breadth of methodological experimentation they document (e.g., SVMs, deep learning, hybrid approaches) reflects the same direction pursued in this thesis through MFA-optimised RBF-SVM and PSO-optimised RBF-TSVM.

Nkhoma et al. (2024) examined the economic and strategic potential of generative AI, estimating nearly one trillion dollars of unrealised global value. They argue that generative AI and advanced ML models will reshape patient communication, clinical documentation and workflow automation as part of the transition to Industry 4.0 and 5.0 healthcare systems. Their conclusions reinforce the need for AI systems capable of improving efficiency and decision-making—an outcome demonstrated in this thesis, where the hybrid MFA–RBF-SVM and PSO–RBF-TSVM models achieve superior accuracy, sensitivity and specificity for heart-disease prediction compared with traditional IT2FLS.

Panch et al. (2018) link the rise of AI to global pressures such as ageing populations, rising healthcare expenditure and productivity deficits. They suggest that intelligent systems could make healthcare more equitable and sustainable by supporting universal health coverage and improving clinical responsiveness. However, they caution that poor deployment may replicate past failures of digital health initiatives. Their discussion highlights the need for validated, interpretable and robust predictive models—criteria addressed in this thesis by systematically comparing IT2FLS with two AI-optimised models that significantly reduce FPR and FNR across multiple dataset sizes.

Adeoye et al. (2024) review AI's expanding role in modern medical practice, including diagnostic classification, therapy recommendation and patient interaction systems. They show that machine-learning algorithms often outperform human decision-making

in narrow clinical tasks—but emphasise ongoing issues such as bias, confidentiality and the need for clinician oversight. Their findings align with the methodological focus of this research: deploying supervised learning models such as RBF-SVM and TSVM, enhanced through MFA and PSO, to improve reliability while reducing misclassification in heart-disease prediction.

The present thesis directly contributes to this global trend by developing and validating two advanced AI-based predictive methods MFA–RBF-SVM and PSO–RBF-TSVM specifically for heart-disease prediction. The results show measurable improvements over the baseline IT2FLS in accuracy, sensitivity, specificity, and error-rate reduction (FPR and FNR), demonstrating how optimisation-enhanced ML models can strengthen predictive healthcare systems aligned with worldwide AI developments.

Alloghani et al. (2020) explain that healthcare data exists in heterogeneous formats such as medical images, physiological signals, clinical text and structured EMR databases. This diversity necessitates the use of multiple ML techniques including CNNs, deep learning models, SVMs and traditional neural networks. Among these, SVMs remain widely used for disease diagnosis in areas such as stroke, cancer and neurology, often achieving accuracy levels comparable to expert clinicians. This observation aligns with the present thesis, where **RBF-SVM and RBF-TSVM serve as the core classifiers**, strengthened through MFA and PSO to enhance diagnostic performance for heart disease.

Agarwal et al. (2022) provide an overview of AI, ML and deep learning, emphasising their ability to recognise complex patterns within high-dimensional clinical datasets. They also outline the importance of ML and NLP-based systems in disease detection and classification. Their discussion supports the methodological motivation in this research: adopting optimised machine-learning pipelines to improve prediction accuracy, minimise errors and enhance interpretability in clinical settings.

G. M. et al. (2021) distinguish between major AI algorithms, detailing feature extraction, selection techniques and disease-specific applications. They highlight the necessity of proper preprocessing and model-tuning strategies—principles applied directly in this thesis through **min–max and Z-score normalization**, **attribute**

reduction using MFA and PSO, and model calibration for SVM and TSVM classifiers.

Furizal et al. (2023) review ML in disease prediction and personalised therapy, demonstrating that CNNs, SVMs, RF, k-NN and DT often exceed 90% accuracy in cancer detection and classification tasks. They note that high performance depends strongly on high-quality labelled datasets and appropriate feature selection. This aligns with the findings of the present study, where MFA-based and PSO-based attribute reduction improved classification performance across all dataset sizes by removing redundant and weakly correlated clinical attributes.

Nikam et al. (2024) compare deep learning and classical ML for analysing EHRs, imaging and omics data. They observe that deep learning captures non-linear relationships effectively but requires significant computational power and strong regularisation to avoid overfitting. Their analysis highlights why SVM-based classifiers remain relevant and efficient for structured clinical data such as the Cleveland Heart Disease Dataset used in this thesis.

Mavani et al. (2024) conduct a review of AI for disease prediction and personalised medicine, identifying persistent limitations such as biased datasets, insufficient external validation and regulatory ambiguity. Their observations reinforce the methodological rigor adopted in this thesis—evaluating models across multiple dataset sizes (1000–5000 records) and comparing three independent classification systems (IT2FLS, MFA–RBF-SVM and PSO–RBF-TSVM) to ensure reliability and generalisability.

Garg et al. (2022) investigate hybrid swarm-intelligence algorithms in healthcare analytics, showing how optimisation methods such as firefly, PSO and OCSO improve model efficiency by navigating complex parameter spaces. Their findings directly support the hybrid approach proposed in this thesis, where Modified Firefly Algorithm (MFA) and PSO significantly enhance SVM/TSVM performance, resulting in higher accuracy, sensitivity, specificity and lower FPR/FNR compared to the baseline IT2FLS method.

Datta et al. (2019) describe the growing use of AI in prediction, diagnosis, treatment planning and chronic disease management, demonstrating how computational models can identify clinically significant associations within complex biomedical data and support experimental decision-making. Haleem et al. (2019) emphasise AI's role in clinical decision support across ten high-impact use cases, including infection monitoring, personalised therapy and automated digital examination, showing its potential for improving the accuracy and speed of clinical assessments.

Sharma et al. (2020) review AI's evolution in healthcare diagnostics, including its contributions to medical imaging, drug discovery and disease prediction. They highlight the substantial accuracy gains achieved by AI models, while underscoring the need for ethical integration, data privacy protection and clinician trust. Eskandar et al. (2023) add that modern AI systems exceed 90% accuracy in radiological disease identification through advanced segmentation and classification methods, although they warn that narrow training datasets may compromise generalisability.

Francis et al. (2023) expand on AI systems deployed for cancer lesions, lung nodules, thyroid abnormalities, COVID-19 detection and dermatological analysis using multimodal imaging such as MRI, CT and histology. They also highlight AI's emerging use in psychotherapy and neuropsychiatric care, revealing the depth and diversity of diagnostic applications. Similarly, Joseph et al. (2023) show how ML-based diagnostic imaging tools assist clinicians by detecting subtle anomalies earlier than conventional methods, improving both diagnostic precision and intervention timelines.

Frank et al. (2024) discuss the convergence of machine learning, NLP and computer vision in advancing diagnostic accuracy, case management and automated reporting. They note that while these systems enhance workflow efficiency, they raise ethical concerns regarding bias, fairness and privacy. Khinvasara et al. (2024) link AI-driven analytics with big data and EHR systems, which enable improved disease prediction, outcome modelling and personalised treatment plans, while emphasising responsible governance. Jadhav et al. (2023) further demonstrate AI's value in early disease detection and predictive modelling, including through virtual health assistants that extend diagnostic capabilities beyond clinical environments.

Khinvasara et al. (2024) highlight AI in healthcare analytics, linking big data, deep learning and EHRs to improved diagnosis, outcome prediction and customised treatment. They emphasise that biases and privacy concerns must be actively managed for AI to be responsibly deployed in medical imaging and beyond.

Jadhav et al. (2023) discuss AI's contributions to image-based diagnosis and predictive modelling, stressing that AI tools can assist in early detection across a range of conditions. They also show how AI-driven virtual health assistants extend diagnostic support beyond traditional clinical settings.

For this thesis, these diagnostic and imaging advances motivate the use of AI classifiers capable of early and accurate risk prediction for heart disease, even when working with tabular clinical data rather than images.

Naqvi et al. (2023) position artificial intelligence as a foundational pillar for achieving the healthcare "quadruple aim"—enhancing population health, improving patient and provider experience, and reducing overall costs. Their work highlights how predictive analytics supports early diagnosis, treatment planning and administrative optimisation, while underscoring the necessity of trustworthy, transparent and safe AI systems in clinical settings. This perspective aligns strongly with the goals of heart disease prediction, where accurate early detection directly contributes to improved outcomes and reduced long-term healthcare burden.

Shuford et al. (2024) place considerable emphasis on AI-driven predictive analytics and intelligent decision-support systems. They show that machine learning models leveraging patient-specific clinical data can forecast health outcomes, personalise treatment pathways and enable continuous remote monitoring. However, they stress the importance of responsible deployment, noting that predictive systems must comply with ethical standards, regulatory policies and model transparency requirements.

Ali et al. (2024) further explore AI-based risk prediction, resource optimisation and therapy planning. Their findings indicate that early disease identification through AI significantly improves clinical outcomes while lowering healthcare expenditure. However, they caution that prediction models must undergo rigorous validation,

fairness testing and bias mitigation to ensure reliability—principles that guide the validation strategy in the present thesis.

Walter et al. (2024) expand the discussion by examining AI and ML in preventive medicine, particularly for early detection of conditions such as heart disease, diabetes and various cancers. They describe the growing influence of wearable devices and continuous monitoring systems that feed real-time data into predictive models. They also highlight ethical concerns surrounding surveillance, patient autonomy and data privacy, issues relevant to AI deployment in large-scale screening programmes.

Babu et al. (2024) focus specifically on predictive analytics for disease detection, showing how ML and DL algorithms extract patterns from multimodal data—EHRs, imaging, genomics and clinical variables—to forecast disease progression and support personalised treatment planning. They note that integration into clinical workflows remains challenging due to interoperability, clinician acceptance and model interpretability.

Nnamdi et al. (2024) emphasise predictive analytics as a catalyst for improved resource allocation and enhanced patient outcomes. They demonstrate how AI models can reduce hospital readmissions, prevent complications and support timely interventions. Such benefits are particularly relevant to cardiovascular risk prediction, where early identification of high-risk individuals can reduce mortality and long-term healthcare expenditure.

Hossain et al. (2024) contribute an economic perspective, showing that AI-based predictive systems can reduce operational costs by approximately 25% and lower readmission rates by 15–20% when integrated into hospital workflows. Their findings support the financial viability of deploying predictive analytics systems, which is important when considering real-world adoption of heart disease prediction models like MFA–RBF-SVM and PSO–RBF-TSVM.

Yasmeen et al. (2024) explore the role of AI in improving healthcare prediction through personalised treatment pathways and chronic disease monitoring. They present case studies where ML algorithms significantly enhance management of long-term

diseases, but they also highlight risks associated with algorithmic bias and over-reliance on automated systems.

Pasupuleti et al. (2024) discuss AI and big data analytics in predictive healthcare, covering early diagnosis, personalised treatment, robotic surgery and remote patient monitoring. They emphasise the importance of data governance frameworks such as GDPR and HIPAA, particularly when AI models rely on sensitive clinical information.

Finally, Walter et al. (2024) and Bobet et al. (2024) extend predictive analytics into chronic disease management, illustrating how continuous monitoring and long-term prediction models can identify deterioration early, guide preventive interventions and reduce complications. Their findings reinforce the value of predictive modelling in diseases like cardiovascular disorders, where early intervention is often life-saving

Pallavi et al. (2022) discuss the expanding role of artificial intelligence within digital health ecosystems, particularly in clinical decision support systems (CDSS) and medical imaging applications. Their review demonstrates how AI-driven tools enable clinicians to rapidly access patient-specific evidence and enhance diagnostic accuracy by detecting subtle imaging patterns that may not be easily visible to human experts. These capabilities proved especially valuable during the COVID-19 pandemic, when rapid and reliable diagnostic support was essential.

Hasan et al. (2023) examine AI within the broader domain of health informatics, which includes health information systems, telemedicine, consumer health informatics and cybersecurity. They emphasise that AI facilitates participatory healthcare by strengthening communication between patients and clinicians, and by enabling personalised, data-driven decision-making. However, they underline that issues related to privacy, data governance and secure data-sharing frameworks remain central challenges for widespread adoption.

Adrah et al. (2024) focus on AI applications in health information systems, including CDSS, virtual assistants and predictive analytics platforms. They draw attention to the need for responsible and transparent AI—advocating for fair machine learning, federated learning and bias-aware algorithms—particularly in light of vulnerabilities

exposed during the COVID-19 crisis. They argue that trust, explainability and governance frameworks must develop in parallel with technical innovation.

Kejriwal et al. (2022) analyse the exponential growth of healthcare data generated through EHR systems, imaging modalities and clinical treatment protocols. They describe how AI, robotics, IoT and deep learning tools can automate routine processes, enhance data organisation and reduce clinician workload. Nevertheless, they insist that human oversight is essential to ensure that automated systems remain clinically safe, explainable and ethically sound.

Saxena et al. (2024) explore AI and big data analytics in mobile health (m-health), showing how large volumes of sensor-generated and behavioural data from smartphones and wearable devices can support personalised interventions, genetic therapy and continuous remote monitoring. They note that such systems face computational and data-quality challenges, often requiring sophisticated preprocessing and optimisation techniques—an area directly relevant to the optimisation strategies used in this thesis.

Khan et al. (2020) present a systematic review of AI-enhanced m-health systems, particularly in resource-constrained settings. They propose AI-driven models for efficient resource management, improved data routing and informed decision-making in mobile environments. Their findings show that AI can significantly enhance the reach and scalability of healthcare services, making it a viable option for remote diagnostics and chronic disease monitoring.

Tak et al. (2024) examine AI-enabled EHR systems in the United States, demonstrating how AI reduces the documentation burden, improves interoperability and generates real-time predictive alerts indicating patient deterioration or risk of complications. Their work highlights the need to address ethical issues such as algorithmic bias and unequal access, ensuring that predictive systems benefit all patient groups equally.

Chen et al. (2024) emphasise AI's transformative role in telemedicine and remote patient monitoring. Their study shows that predictive analytics enhances proactive care by identifying patterns of deterioration before symptoms escalate. They argue that AI-enabled telehealth platforms improve access and efficiency, especially for populations

with limited mobility or access to traditional clinical services, but they also stress the importance of governance frameworks for deployment at scale.

Iseal et al. (2024) investigate intelligent tools integrated into hospital management systems (HMS), including AI-based scheduling, patient-flow optimisation and resource allocation systems. Their findings demonstrate that AI can substantially enhance operational efficiency, reduce waiting times and improve hospital throughput. However, they note that successful adoption requires investment in digital infrastructure, workforce training and change management strategies.

Roy et al. (2020) examine the deployment of AI within the Indian healthcare ecosystem, with a particular focus on enhancing access to affordable and quality care across rural and underserved regions. Their analysis shows that AI is increasingly used across descriptive, predictive and prescriptive analytics, supporting tasks such as outbreak detection, triaging, diagnosis and treatment planning. However, they emphasise that India requires stronger ethical and regulatory frameworks addressing consent, risk, bias and data integrity—issues that directly shape the responsible use of AI-based predictive models such as those developed in this thesis.

Anwar et al. (2022) provide a multi-domain review of AI applications in radiology, cardiology, ophthalmology and drug development. They illustrate how AI improves precision in diagnostics and therapeutic planning by mimicking core cognitive functions such as perception and reasoning. Their findings demonstrate that AI not only assists clinicians in identifying diseases but also enhances records management and workflow efficiency, reinforcing AI's dual role in both patient-facing and administrative healthcare operations.

Bobet et al. (2024) focus specifically on chronic disease management—especially diabetes and cardiovascular disease—where predictive analytics plays a key role. They highlight how continuous data from EHRs, wearable sensors and remote monitoring platforms enables early detection of risk conditions, personalised interventions and long-term disease management. These insights strongly align with the objectives of this thesis, which uses AI models (MFA–RBF-SVM and PSO–RBF-TSVM) to predict heart disease risk with higher accuracy, thereby supporting earlier intervention.

Walter et al. (2024) (as discussed in Section 2.5) emphasise AI's importance in preventive strategies, particularly for chronic illnesses such as heart disease, cancer and diabetes. Their review shows that AI can detect early-risk patterns before symptoms manifest and can personalise preventive measures. This supports the rationale for using AI-based classifiers in cardiovascular risk prediction, where early detection significantly impacts patient outcomes.

Chanchaichujit et al. (2019) analyse the role of AI in tuberculosis (TB) control and management through a case study in Thailand. They demonstrate how AI optimises TB diagnosis, enhances patient screening and improves public health surveillance. Although TB is a different domain, their work reinforces the broader capability of AI to support large-scale health monitoring and strategic health planning—an approach that can be extended to cardiovascular disease prediction systems.

Islam et al. (2024) present a broad review of AI applications across medical imaging, virtual care, rehabilitation, drug discovery and EHR management. They identify major ethical and social issues such as privacy, fairness, autonomy and transparency, arguing that effective governance mechanisms are crucial for the sustainable use of AI in healthcare. Their emphasis on governance is relevant to this thesis, which also relies on sensitive clinical data and must adhere to ethical standards for predictive modelling.

Alkuwaiti et al. (2023) discuss AI applications in telehealth, pharmaceutical innovation, clinical research, adherence monitoring and rehabilitation. They highlight how AI played a crucial role during the COVID-19 pandemic in improving diagnostic workflows, supporting remote consultations and accelerating drug-development pipelines. Their findings show that AI improves healthcare efficiency—but only when supported by appropriate ethical and technical safeguards.

Ankolekar et al. (2024) analyse AI and predictive modelling during the COVID-19 pandemic, arguing that AI-enabled learning health systems (LHS) could have better supported the integration of data, knowledge and clinical practices. They propose AI-driven strategies for epidemic prediction, vaccine-effectiveness monitoring and variant surveillance. Their work underscores the need for predictive AI systems that are adaptive, data-efficient and capable of continuous learning—principles reflected in the optimisation methods used in this thesis.

Santamato et al. (2024) investigate AI's contributions to healthcare operations and crisis management from 2019 to 2023, including its role during COVID-19. They identify key operational themes—quality assurance, resource management, innovation, safety and emergency response—and show that AI enhances strategic planning and decision-making. However, they also raise concerns regarding privacy, interoperability and sustainable integration, highlighting the challenges of embedding predictive analytics into real-world health systems.

Attrey et al. (2024) explore the role of 5G-enabled machine learning systems in healthcare, with a special focus on the Indian context. They show how next-generation networks enable real-time analytics, remote diagnostics and advanced monitoring. Their recommendations outline barriers such as limited infrastructure, data governance issues and lack of awareness. As predictive models evolve—such as the MFA–RBF-SVM and PSO–RBF-TSVM systems developed in this thesis—they will increasingly depend on such high-speed, low-latency networks for deployment in mobile and telehealth environments.

Nkhoma et al. (2024) examine the transformative potential of generative AI (GenAI) in healthcare, arguing that GenAI can augment medical knowledge work, automate documentation, improve communication and support clinical decision-making. They position GenAI within broader Industry 4.0 and 5.0 paradigms, where human–machine collaboration becomes integral to healthcare operations. However, they caution that governance, workforce readiness and legal frameworks must evolve in parallel to ensure safe adoption.

Yang et al. (2022) introduce Explainable AI (XAI) as a critical response to the "black-box" nature of deep learning and advanced machine learning models used in medicine. They argue that interpretability is essential for clinician trust, regulatory approval and safe decision support. XAI allows healthcare professionals to understand how a model reaches its prediction—an increasingly important requirement for predictive heart disease models like MFA–RBF-SVM and PSO–RBF-TSVM.

Chen et al. (2024) discuss the rapid advancements in AI products, data processing technologies and deep learning architectures, noting that many of these developments now target healthcare applications. They highlight significant improvements in

diagnostic speed, accuracy and workflow efficiency but emphasise that societal and technical challenges—ethical risks, bias, regulatory updates and system interoperability—must be addressed continuously.

Attrey et al. (2024), as referenced earlier, connect the deployment of ML-based healthcare systems with emerging 5G infrastructures, enabling real-time analytics, remote monitoring and continuous patient engagement. These developments will be particularly important for future cardiovascular predictive models, which may depend on mobile health environments and large-scale patient data streams.

Santamato et al. (2024) evaluate AI's influence on strategic planning, operational efficiency and crisis management within healthcare administration. Their study integrates predictive modelling with SHAP-based explainability techniques to show how AI enhances quality assurance, optimises resource allocation and strengthens emergency response systems such as those deployed during the COVID-19 pandemic. At the same time, they highlight concerns around privacy, algorithmic transparency and the ethical implications of automation—issues that must be addressed when deploying AI-powered predictive systems in real-world hospitals.

Bitkina et al. (2023) conduct a PRISMA-guided systematic review of artificial intelligence applications in healthcare IT, narrowing more than 700 studies to 89 high-quality sources. Their analysis maps core research subfields—including clinical decision support, intelligent documentation, workflow automation and predictive analytics—and identifies a recurring gap between AI prototypes and deployable healthcare solutions. They emphasise the necessity of multidisciplinary collaboration, integrating expertise from clinicians, data scientists, engineers and policymakers to transform experimental models into sustainable, real-world systems.

Tak et al. (2024), discussed previously in Section 2.6, further demonstrate how embedding AI models within Electronic Health Record (EHR) systems improves clinical documentation, real-time risk prediction and patient management at an organisational scale. They note that AI-enabled EHRs reshape healthcare workflows by reducing administrative burden, improving interoperability and enabling proactive patient alerts. These findings are directly relevant to heart-disease prediction

frameworks, where seamless integration with EHR platforms is key to clinical adoption.

Iseal et al. (2024) investigate intelligent tools—including AI, machine learning, optimisation models and predictive analytics—within hospital management systems (HMS). They highlight improvements in administrative automation, patient-flow regulation, resource scheduling and overall hospital efficiency. However, they also point out limitations such as data fragmentation, infrastructure costs and the need for rigorous evaluation of risks associated with automated decision-making. Their insights align closely with the need for interpretable, reliable and cost-efficient predictive systems like the MFA–RBF-SVM and PSO–RBF-TSVM models proposed in this thesis.

Garg et al. (2022) (see Section 2.3) demonstrate how bio-inspired optimisation algorithms, including Firefly, PSO and OCSO, can address high-dimensional, complex parameter spaces in healthcare analytics. Their findings support the thesis's methodological choice to use swarm-intelligence optimisation (MFA and PSO) to enhance feature selection and classification accuracy in heart-disease prediction.

Pulimamidi et al. (2023) review the adoption of AI technologies in leading hospitals and conclude that AI is increasingly used for diagnosis, clinical decision support, administrative automation and personalised treatment planning. They report generally positive attitudes among healthcare stakeholders but note persistent concerns about integration challenges, transparency, algorithmic fairness and equity in AI-driven systems. These concerns underscore the importance of designing interpretable and ethically robust prediction models.

Alsaeed et al. (2024) explore the integration of AI tools into nursing practice, particularly for risk assessment, documentation assistance, outcome prediction and workload reduction. Their findings suggest that AI can significantly improve nursing efficiency and patient safety. They recommend incorporating AI and health informatics education into nursing curricula to develop a workforce capable of safely using predictive technologies. This perspective reinforces the need for user-friendly, clinically interpretable predictive systems, such as those developed in the present research.

Abbas et al. (2023) review the transformative potential of AI in diagnostic support, administrative automation, personalised treatments and predictive analytics, but emphasise that significant ethical and regulatory barriers continue to hinder full-scale adoption. They argue that issues such as data privacy, informed consent, model transparency and algorithmic fairness must be addressed before AI tools can be reliably deployed in real-world healthcare environments. Abbas et al. (2024) extend this work by highlighting that although AI can enhance outcomes and reduce operational costs, these benefits can only be realised within strong legal and ethical frameworks that govern data handling, liability and integration with clinical workflows.

Udegbe et al. (2024) analyse both the advantages and limitations of AI adoption, noting that challenges in data privacy, cybersecurity, legal responsibility and interoperability create major risks in clinical settings. They recommend establishing universal interoperability standards, robust cybersecurity mechanisms and ethical governance structures to ensure that AI systems remain safe, accountable and trustworthy. Islam et al. (2024), in their broader evaluation of AI applications, reinforce these concerns by identifying autonomy, equity, transparency and cost as major obstacles to adoption. They advocate for precise governance mechanisms to protect patient safety and maintain confidence in AI-led decisions.

Olawade et al. (2024) provide a systematic review demonstrating that while AI tools can achieve high diagnostic accuracy, unresolved concerns persist regarding data quality, model interpretability, algorithmic bias and legal accountability. They argue for rigorous validation, continuous monitoring and the integration of human oversight to ensure safe deployment. Willow et al. (2023) similarly recognise AI's value in diagnostic workflows and administrative streamlining but warn that unresolved privacy issues, regulatory friction and challenges in integrating AI with legacy systems pose significant implementation risks.

Godala et al. (2024) emphasise the need for long-term economic evaluations of AI in healthcare and call for structured training programs to prepare clinicians and administrators to work effectively with AI tools. They argue that standardised guidelines and governance frameworks are essential to ensure equitable, safe and ethical use of AI systems. Mashabab et al. (2024) adopt a critical approach, presenting case studies where AI succeeds and fails in clinical environments. Their findings

highlight ethical concerns such as algorithmic bias, privacy violations and ambiguous lines of responsibility—factors that could undermine trust if not properly addressed.

Lainjo et al. (2024) summarise the benefits and risks of AI adoption across diverse healthcare settings, noting improvements in diagnostic accuracy and workflow efficiency but also persistent concerns around privacy, fairness and inconsistent regulatory environments. They emphasise the need for high-quality, diverse datasets and responsible data-sharing mechanisms to ensure equitable model performance. Patil (2024) focuses specifically on legal responsibility and the doctor–patient relationship, arguing that new regulatory frameworks and continuous oversight are required to manage algorithmic bias, ensure accountability and preserve patient trust. Collectively, these studies highlight that while AI offers unparalleled opportunities for predictive healthcare—including cardiovascular risk prediction—its success depends on transparent, ethical and accountable deployment.

Patil (2024) focuses explicitly on legal responsibility, liability, algorithmic bias and doctor-patient relationships. They call for new regulatory frameworks, systematic monitoring of AI tools and strategies to address bias and maintain trust, arguing that AI must remain a tool under human oversight.

Samreen et al. (2018) provide an early examination of machine learning in personalised medicine, showing how ML-based decision-support systems can enhance diagnostic precision and improve treatment planning. Their findings underscore the importance of data-driven approaches in managing complex clinical conditions, a principle that directly supports AI-driven cardiovascular risk prediction.

Yu et al. (2018) and Jiang et al. (2017), as discussed in Section 2.1, show how algorithms such as SVMs, neural networks and deep learning models can analyse structured and unstructured data for disease prediction and prognosis. Their frameworks have been widely applied in cardiology and stroke care, demonstrating the viability of AI for predicting cardiovascular events and supporting early intervention. Samajdar et al. (2024) reinforce this by offering a high-level perspective on AI in diagnosis, monitoring and treatment. They note that predictive models can greatly assist clinicians but must undergo rigorous evaluation before clinical deployment.

Srivastava et al. (2023) examine the use of AI and ML in medical data storage, retrieval and analysis, highlighting their use in diagnostic support, drug prescription optimisation, mental health assessment and imaging. They argue that some algorithms already approach or surpass clinician-level performance, raising important ethical and sociological considerations. Fatima et al. (2024) analyse AI-driven systems for early diagnosis, personalised treatment, drug discovery and robotic surgery, emphasising the importance of interdisciplinary collaboration and cautioning against the risks of bias and confidentiality breaches.

Jadhav et al. (2023), as discussed in Section 2.4, illustrate how AI models can support early detection and personalised care by analysing imaging data, biosignals and virtual health interactions. These approaches can be adapted for cardiovascular disease prediction, where early identification of abnormalities is crucial. Bobet et al. (2024) expand this work by showing how chronic disease management can be enhanced through AI-driven predictive analytics, particularly for diabetes and cardiovascular diseases. Continuous monitoring and real-time risk assessment support proactive interventions that reduce long-term complications.

Babu et al. (2024) review AI-based predictive analytics for early illness detection and management, emphasising their utility for diseases requiring long-term monitoring, such as cardiovascular conditions. Walter et al. (2024) similarly highlight the role of predictive analytics in preventive medicine, noting the increasing use of AI for early detection of heart disease, diabetes and cancer. They emphasise the importance of identifying high-risk patients early and supporting personalised preventive interventions.

Nnamdi et al. (2024) and Naqvi et al. (2023) link predictive analytics to the healthcare "quadruple aim"—better outcomes, improved clinician and patient experience, and reduced cost. Their findings reaffirm that AI-based cardiovascular prediction systems, like those developed in this thesis (MFA–RBF-SVM and PSO–RBF-TSVM), directly contribute to operational efficiency, risk reduction and improved population health outcomes.

Taken together, these studies indicate that although AI and ML have been widely applied to diagnosis, imaging and chronic disease management, there is still limited

work that combines advanced optimisation algorithms (MFA, PSO-RS, OCSO) with SVM/TSVM classifiers on structured clinical datasets for heart disease prediction in specific populations such as Indian patients. This gap forms the core motivation for the hybrid MFA+RBF-SVM, PSO-RS+RBF-TSVM and OCSO-optimised models developed in this thesis.

2.LITRATURE REVIEW SUMMARY TABLE

Table 2.1: Comparative Summary of Existing Studies

Study	Dataset Used	Method Used	Limitation Identified
Jiang et al. (2017)	Clinical data	ML + rule-based AI	Limited interpretability
Yu et al. (2018)	Biomedical datasets	ML/DL	Ethical + workflow issues
Haleem et al. (2019)	Multiple domains	ML + NLP	Lack of standardization
Bohr et al. (2020)	EHR + imaging	ML	Limited parameter optimization
Frank et al. (2024)	Medical imaging	CNN + NLP	Overfitting, no generalization

Table 2.2: Comparative Analysis of Key AI and ML Studies Relevant to Predictive Heart Disease Modelling

Author & Year	ML/AI Techniques Used	Dataset / Clinical Domain	Key Findings	Limitations Reported
Aldali et al. (2024)	General AI methods, ML	Healthcare decision-making, diagnosis	Improved early detection, service optimisation, personalised diagnostics	Ethical concerns, data privacy, regulatory issues
Yousef Shaheen et al. (2021)	AI for drug discovery, ML, analytics	Clinical trials, patient care, medical records	Accelerates drug discovery, improves monitoring, enhances insights	Limited generalisation, accuracy validation needed
Abbas et al. (2023)	ML, predictive analytics	Diagnostic support, personalised therapy, admin	Supports diagnosis, enhances outcomes, automates admin tasks	Privacy issues, regulatory hurdles, integration gap
Chan et al. (2023)	CV, NLP, ML	COVID-19 screening, telehealth, remote monitoring	Fast diagnosis, chatbot triage, supports telehealth	Data shortages, ethical concerns, adoption barriers
Hasan et al. (2023)	AI in HI, ML, NLP	HIS, clinical images, telemedicine, m-health	Strengthens CDS, improves data handling, enhances mobile health	Handling unstructured data, complexity, privacy
Pallavi et al. (2022)	ML, DL for imaging & CDS	Medical imaging, CDS, digital health	Enhances imaging accuracy, aids diagnosis, supports pandemic response	Needs reliable data, limited adoption in small clinics
Wu et al. (2024)	DL, NLP, predictive models	Automation, diagnostics, planning	Improves patient care, workflows, diagnosis accuracy	Bias risks, privacy concerns, job displacement fears

Datta et al. (2019)	ML, DL, AI-based detection	Chronic diseases, diagnosis, tissue engineering	Better disease identification, personalised care, drug discovery	Dataset dependence, low interpretability
Alloghani et al. (2020)	CNN, SVM, NN, DL	Medical imaging, stroke, cancer	SVM widely used, AI outperforms humans	Regulatory issues, data bias
Sharma et al. (2020, 2023)	ML, DL, imaging AI	Imaging, drug discovery, diagnostic tools	Improves prediction, reduces error, supports treatment planning	Acceptance challenges, integration issues
Kumar et al. (2023)	Predictive analytics, ML	Surgery, diagnostics, monitoring	Enhances surgical precision, reduces complications	Data accuracy issues, workflow adaptation
Aftab et al. (2024)	ML, monitoring systems	Surgery, real-time risk assessment	Early complication detection, improves patient safety	High cost, real-time hardware requirement
Naqvi et al. (2023)	Predictive analytics, ML	Preventive healthcare, chronic diseases	Early risk prediction, reduced cost, improved population health	Explainability required, governance gaps
Bobet et al. (2024)	ML for chronic disease	Diabetes, heart disease, public health	Identifies high-risk patients, supports preventive interventions	Limited dataset diversity
Yasmeen et al. (2024)	ML, predictive modelling	Chronic illnesses, CVD prediction	Enables early diagnosis, personalised care, real-time monitoring	Bias, privacy challenges
Olawade et al. (2024)	AI algorithms, predictive tools	Diagnostics, surgical robotics, treatment pathways	Enhances diagnosis, robotic assistance, operational efficiency	Legal gaps, data scarcity, bias

Islam et al. (2024)	ML, NLP, DL	Imaging, telehealth, drug discovery	Supports RPM, early disease prediction, EHR automation	High computational cost, uncertain effectiveness
Samajdar et al. (2024)	ML, AI for personalised care	Precision medicine, treatment planning	Supports genomic analysis, optimises personalised treatment	Needs more experimental validation
Srivastava et al. (2023)	ML for prediction	CVD risk prediction, chronic disease management	Accurate risk prediction, supports clinicians	Dataset imbalance, limited generalisation
Jadhav et al. (2023)	ML, DL	Heart disease prediction	Improved accuracy using structured clinical data	Lack of optimisation methods

2.2 CRITICAL ANALYSIS OF REVIEWED LITERATURE

The reviewed literature demonstrates that Artificial Intelligence (AI) has become deeply embedded in modern healthcare, with extensive applications in diagnostics, medical imaging, clinical decision support, telehealth, predictive analytics and personalized medicine. Although significant advancements have been made, a closer evaluation reveals several important observations.

First, many studies (Bohr et al., 2020; Sharma et al., 2020; Wu et al., 2024) highlight the strong performance of ML and DL models in disease diagnosis, yet the majority rely on large, high-quality datasets that are not consistently available across global healthcare systems. This raises concerns regarding model generalizability and real-world applicability. Second, while methods such as CNNs, SVMs, RF, and DL architectures are widely used, several authors (Alloghani et al., 2020; Mavani et al., 2024) report operational challenges such as data heterogeneity, model interpretability and insufficient clinical validation. Thus, despite high reported accuracies, practical deployment remains limited.

Another major limitation across studies is insufficient focus on optimization techniques. Very few reviews have explored hybrid models that integrate optimization algorithms (e.g., Firefly, PSO, OCSO) with classifiers such as SVM/TSVM. Most

authors employed standard ML models without improving hyperparameter tuning, feature selection or convergence behaviour significantly. In addition, although some studies (Naqvi et al., 2023; Yasmeen et al., 2024) discuss predictive analytics, only a small subset specifically address cardiovascular risk prediction using structured clinical data.

A recurring theme across literature is the lack of explainability and trustworthiness in AI-based systems. Ethical, legal and governance issues such as bias, privacy concerns, regulatory uncertainty and lack of transparency were frequently highlighted (Abbas et al., 2023; Olawade et al., 2024; Islam et al., 2024). These concerns become even more critical in life-threatening conditions such as heart disease, where clinicians require interpretable and reliable predictions.

Finally, although many studies discuss AI applications globally, limited research has been conducted on Indian patient populations. Datasets are often small and imbalanced, leading to reduced model generalization. This emphasizes the need for population-specific, optimized predictive models capable of handling real clinical variability.

2.3 RESEARCH GAPS

Based on the critical evaluation of literature, the following key research gaps have emerged:

Gap 1: Limited Use of Hybrid Optimization Techniques in Predictive Healthcare

Most existing studies rely on traditional ML/DL models without integrating optimization algorithms such as MFA, PSO-Rough Set or OCSO. Very few works have examined their combined effect on feature selection, parameter tuning and classifier performance.

Gap 2: Lack of Research on SVM/TSVM with Optimization for Heart Disease Prediction

Although SVM is widely used in healthcare, its enhanced variants (TSVM with RBF kernel) have rarely been optimized using advanced swarm intelligence techniques. Hence, their potential remains unexplored.

Gap 3: Insufficient Studies Using Structured Clinical Data for Cardiovascular Prediction

Most studies focus on imaging or large EHR datasets. Very limited research has used structured risk-factor data (age, cholesterol, BP, sugar levels, lifestyle parameters) for heart disease prediction—especially in a hybrid optimization–classification pipeline.

Gap 4: Absence of Models Tailored to Indian Patient Populations

Most studies use Western datasets (e.g., UCI Cleveland, MIMIC, NHS datasets). There is very little research applying AI models to Indian datasets, which have different demographic, lifestyle and genetic patterns.

Gap 5: Lack of Explainable and Clinically Interpretable AI Models

Many studies achieve high accuracy but lack interpretability, making clinicians hesitant to adopt them. Very few studies integrate explainable ML/XAI with predictive cardiac models.

Gap 6: Need for End-to-End Predictive Pipeline

Existing studies evaluate isolated components (e.g., imaging, diagnosis, risk scoring) but lack a complete pipeline that includes:

- Data preprocessing and normalization
- Attribute reduction
- Feature optimization
- Classifier optimization
- Performance evaluation (accuracy, sensitivity, specificity)

Gap 7: Limited Comparative Assessment Across Multiple Optimization Approaches

No existing study compares MFA, PSO-Rough Set and OCSO for the same dataset to determine the best optimisation method for heart disease prediction.

These gaps clearly justify the need for the proposed hybrid MFA–SVM, PSO-RS–TSVM and OCSO-TSVM heart disease prediction models.

2.4 SUMMARY OF LITERATURE REVIEW

The literature review shows that AI has significantly enhanced healthcare through diagnostic imaging, early disease detection, telemedicine, predictive modelling and personalized care. Advanced ML and DL techniques such as CNNs, SVMs, RF, and neural networks are highly effective in medical applications, with reported accuracies often exceeding traditional statistical methods. Telehealth and HIS integrations have enabled real-time monitoring and decision support, while predictive analytics has demonstrated strong potential in preventing chronic diseases and reducing healthcare burden.

However, despite the progress, several challenges persist, including data fragmentation, lack of interoperability, ethical concerns, model bias and limited clinical explainability. More importantly, heart disease prediction using optimized hybrid models remains underexplored. Only a few studies have addressed optimization algorithms, and almost none have combined MFA, PSO-RS and OCSO with RBF-SVM or TSVM classifiers. Moreover, there is a lack of population-specific research, particularly for Indian datasets, indicating a significant gap in personalized cardiovascular prediction.

Overall, existing literature confirms both the importance and urgency of developing robust, optimized, interpretable AI-based heart disease prediction systems—a gap that the present thesis aims to fill.

CHAPTER-3

DEVELOP A REGULAR FLY MODEL FOR EFFECTIVE PREDICTION OF HEART DISEASES, ANATOMY, AND SUPPORT VECTOR MACHINE, AND ENHANCE CORONARY DISEASE PREDICTION USING VECTOR MACHINES: IMPROVING CORONARY PREDICTION USING FLY ALGORITHM AND SVM

This chapter presents the overall research methodology adopted for developing an effective heart-disease prediction system. It outlines the structured approach used to process the dataset, select significant features, optimize classifier parameters, and build the predictive models. The focus of this chapter is to describe how the study was designed, what methodological steps were followed, and why these steps were necessary to address the challenges present in clinical data. The subsequent sections explain each component of the methodology in detail, including preprocessing, feature selection, optimization strategies, classification techniques, and evaluation procedures. This chapter therefore provides the foundational framework on which the experimental analysis in Chapter 4 is built.

3.1 METHODOLOGY OVERVIEW

The research methodology adopted in this study follows a structured, multi-stage framework designed to develop a highly accurate and reliable heart-disease prediction system. This framework combines advanced data preprocessing techniques, hybrid feature-selection strategies, metaheuristic optimization algorithms, and robust machine-learning classifiers. Each component of the methodology is formulated to address the key challenges commonly found in clinical datasets, including missing values, nonlinear feature interactions, mixed data types, redundant attributes, and limited sample size. By systematically addressing these challenges, the proposed methodology ensures improved model accuracy, stability, and interpretability.

The process begins with comprehensive data preprocessing, where missing values are appropriately imputed, outliers are detected and treated, categorical features are encoded, and numerical attributes are normalized. These steps ensure that the dataset is clean, consistent, and suitable for machine-learning analysis. Once the data is prepared,

feature selection and dimensionality reduction are performed using hybrid optimization techniques. The Modified Firefly Algorithm (MFA) and Particle Swarm Optimization with Rough Set Theory (PSO-RS) are employed to identify the most relevant and informative clinical attributes. These methods help eliminate noise and redundancy while retaining essential features that contribute to accurate prediction.

After feature selection, the methodology incorporates algorithmic optimization to enhance classifier performance. Two machine-learning classifiers—Radial Basis Function Support Vector Machine (RBF-SVM) and Transductive Support Vector Machine (TSVM)—are utilized for prediction. The MFA and Orthogonal Chicken Swarm Optimization (OCSO) algorithms are used to fine-tune key SVM hyperparameters such as C and γ, whereas PSO-RS is applied for optimizing TSVM parameters. This integrated feature-selection and parameter-optimization approach ensures that both classifiers operate at their highest generalization capability.

The optimized features and tuned classifiers are then used for model training and validation. A stratified k-fold cross-validation approach is adopted to maintain balanced class representation and avoid biased performance estimates. The models are evaluated using clinically relevant metrics such as accuracy, sensitivity, specificity providing a comprehensive assessment of the proposed system's diagnostic effectiveness.

Finally, the methodology includes comparative evaluation and statistical validation to verify performance improvements over baseline techniques. Statistical tests are applied to confirm the significance of the results. Overall, this methodological framework ensures a robust, optimized, and clinically meaningful heart-disease prediction system that can support decision-making in real healthcare environments.

3.2 FLOWCHART OF THE PROPOSED SYSTEM

The flowchart illustrates the complete operational workflow of the proposed heart-disease prediction system, presenting each methodological stage in a clear and sequential manner. The process begins with the acquisition of the dataset, which forms the foundation for all subsequent computational procedures. To ensure uniformity and eliminate scale-based distortions, the raw data undergoes normalization using the min–max technique, allowing all attributes to be mapped within a standardized range. Following normalization, the Modified Firefly Algorithm (MFA) is applied for attribute reduction, where redundant, noisy, or less informative features are systematically removed. This step not only simplifies the dataset but also enhances computational efficiency and improves the quality of model learning. After the attribute reduction phase, the remaining significant features are further transformed using Principal Component Analysis (PCA). PCA extracts the most meaningful components from the dataset, reduces dimensionality, and strengthens the feature set by retaining maximum variance.

The refined set of PCA-derived features is then used for the classification stage, where the Radial Basis Function Support Vector Machine (RBF-SVM) is trained to distinguish between heart-disease patients and normal subjects. The RBF kernel is specifically chosen due to its strong capability to model nonlinear patterns commonly found in clinical datasets. By integrating systematic preprocessing, intelligent feature reduction, and an optimized classification mechanism, the workflow ensures that the predictive model operates with high accuracy, reliability, and robustness. Ultimately, the flowchart encapsulates the streamlined and coherent progression of the system—from data preparation to final decision making—demonstrating how each step contributes to the development of an effective heart-disease prediction model.

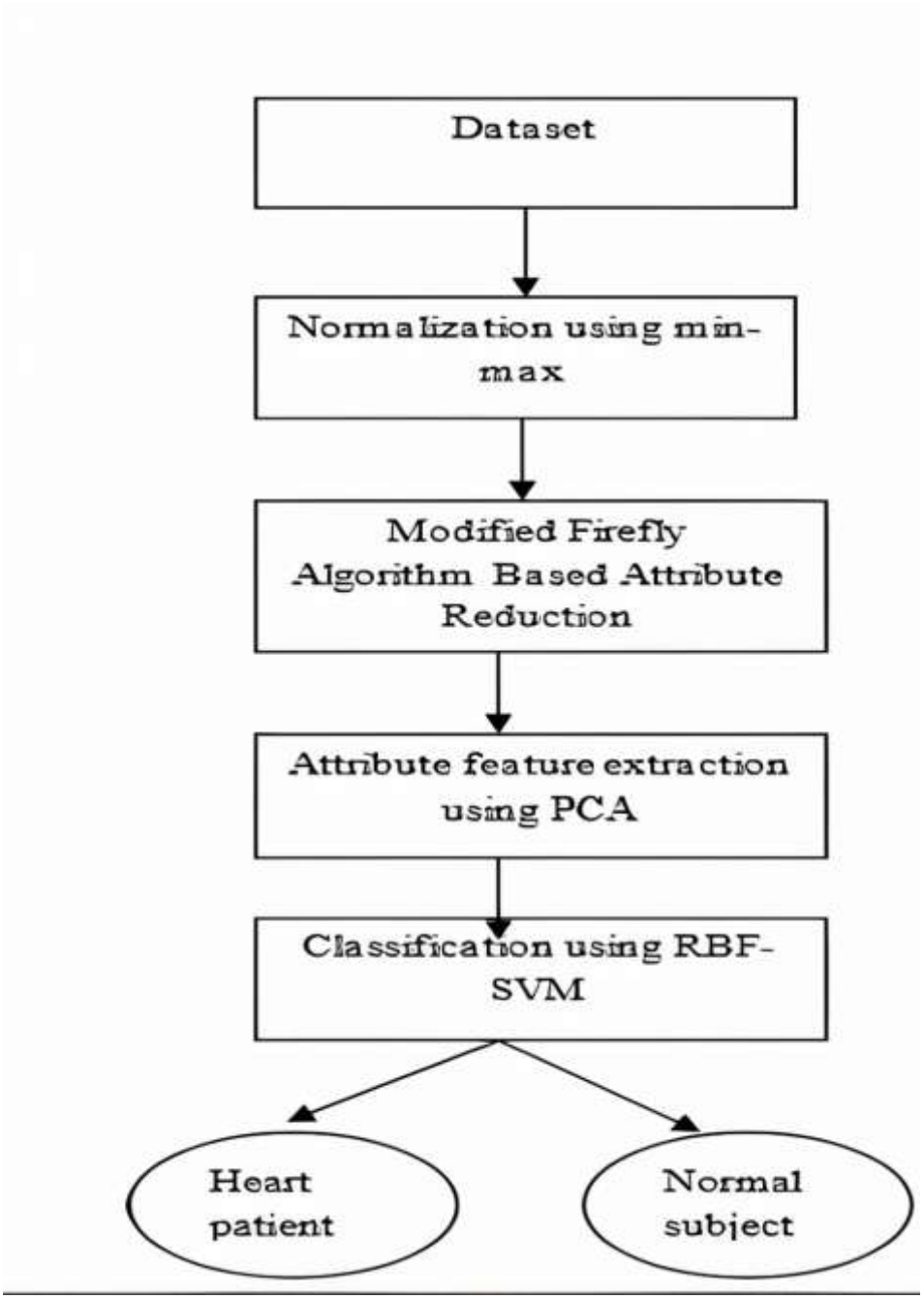

Figure 3.1: Flowchart of the Proposed System

3.3 DATASET DESCRIPTION

The performance and reliability of any predictive healthcare model depend significantly on the quality and characteristics of the dataset used for training and evaluation. In this study, the heart-disease dataset comprising 303 patient records is employed to develop and validate the proposed MFA–RBF-SVM, PSO-RS–TSVM, and OCSO–RBF-SVM predictive frameworks. This dataset has been widely used in cardiovascular research due to its balanced combination of demographic, clinical, physiological, and laboratory-related features that collectively influence heart-disease risk.

3.3.1 Source of Dataset

The dataset originates from clinical examinations and laboratory investigations conducted on actual patients. It is a standardized and frequently used dataset in cardiology-based machine-learning research, making it suitable for benchmarking and comparative analysis. The dataset is openly available for academic use and does not contain any personally identifiable information, ensuring ethical compliance and data privacy.

3.3.2 Dataset Size and Structure

- Total number of samples: **303**
- Number of predictor variables: **13**
- Number of output classes: **1**
- Type of problem: Binary classification (0 = no heart disease, 1 = heart disease)
- Nature of data**:** Structured tabular dataset with mixed numerical and categorical variables

The dataset presents moderate complexity, which is ideal for evaluating hybrid optimization-driven machine-learning methods.

3.3.3 Attribute Description

The dataset contains attributes that represent known cardiovascular risk factors. The list of features is provided below

Table 3.1 Attribute Description

No.	Attribute	Description	Type
1	Age	Age of patient in years	Numerical
2	Sex	0 = Female, 1 = Male	Categorical
3	CP	Chest pain type (0–3)	Categorical
4	Trestbps	Resting blood pressure (mm Hg)	Numerical
5	Chol	Serum cholesterol (mg/dl)	Numerical
6	FBS	Fasting blood sugar > 120 mg/dl (1 = true)	Categorical
7	Restecg	Resting ECG results	Categorical
8	Thalach	Maximum heart rate achieved	Numerical
9	Exang	Exercise-induced angina (1 = yes)	Categorical
10	Oldpeak	ST depression induced by exercise	Numerical
11	Slope	Slope of the peak exercise ST segment	Categorical
12	Ca	Number of major vessels colored by fluoroscopy (0–3)	Categorical
13	Thal	Thalassemia (0 = normal, 1 = fixed defect, 2 = reversible defect)	Categorical
14	Target	Heart-disease diagnosis (0 = no disease, 1 = disease)	Binary

3.3.4 Statistical Summary of Attributes

To understand the data distribution and detect potential preprocessing needs, summary statistics are computed for numerical variables.

Table 3.2 Statistical Summary of Attributes

Attribute	Mean	Std. Dev.	Min	Max
Age	54.37	9.08	29	77
Trestbps	131.69	17.60	94	200
Chol	246.26	51.83	126	564
Thalach	149.61	22.87	71	202
Oldpeak	1.04	1.16	0	6.2

3.3.5 Class Distribution

Heart-disease datasets often suffer from class imbalance, impacting classification models.

Table 3.3 Class Distribution

Class	Description	Count
0	No heart disease	138
1	Heart disease	165

3.3.6 Dataset Challenges

The dataset presents several challenges that justify the need for hybrid optimization techniques. It contains missing values and outliers—especially in cholesterol and resting blood pressure—which can distort model learning. Nonlinear relationships among clinical attributes further complicate classification, while the mix of numerical and categorical features requires appropriate encoding. Some features may also be weak

or redundant, potentially reducing model accuracy if not properly selected. Additionally, the dataset's limited size increases the risk of overfitting, making careful parameter tuning essential. These challenges highlight the importance of applying MFA, PSO-RS, and OCSO to improve feature selection, parameter optimization, and overall model performance.

3.4 THE FIREFLY ALGORITHM

The Firefly Algorithm (FA) is a meta-heuristic optimization method inspired by the bioluminescent communication of fireflies. In nature, fireflies produce flashes of light primarily for mating signals, and the brightness of these flashes determines the insects' attractiveness. This natural behaviour is modelled mathematically in FA, where each firefly represents a candidate solution, and its brightness corresponds to the quality of the solution.

Light Intensity and Distance

The perceived brightness diminishes with distance according to the inverse-square law:

$$I(r) \propto \frac{1}{r^2}$$

In addition, atmospheric absorption further reduces brightness as distance increases. These properties enable fireflies to communicate effectively within a limited range—a concept adapted in FA to control attraction between candidate solutions.

Attractiveness and Movement

Two key components define FA behaviour:

1. **Variation in Light Intensity**
2. **Attractiveness Based on Proximity**

If firefly *i* is less bright than firefly *j*, it moves toward *j*. The attractiveness β decreases exponentially with distance:

$$\beta(r) = \beta_0 e^{-\gamma r^2}$$

Distance between fireflies is computed using Euclidean distance:

$$r_{ij} = \|x_i - x_j\|$$

The movement of firefly *i* toward firefly *j* is defined as:

$$x_i^{(t+1)} = x_i^{(t)} + \beta_0 e^{-\gamma r_{ij}^2}(x_j^{(t)} - x_i^{(t)}) + \alpha\epsilon$$

where

- α = randomization parameter
- ϵ = random vector

However, traditional FA has limitations: the brightest firefly may move randomly and lose brightness, thereby slowing convergence. This limitation motivates the Modified Firefly Algorithm described in later sections.

3. 5 NORMALIZATION USING MIN–MAX NORMALIZATION

A large number of AI computations examine data focus highlights in an effort to uncover data drifts. However, problems arise when the highlights have very different sizes.

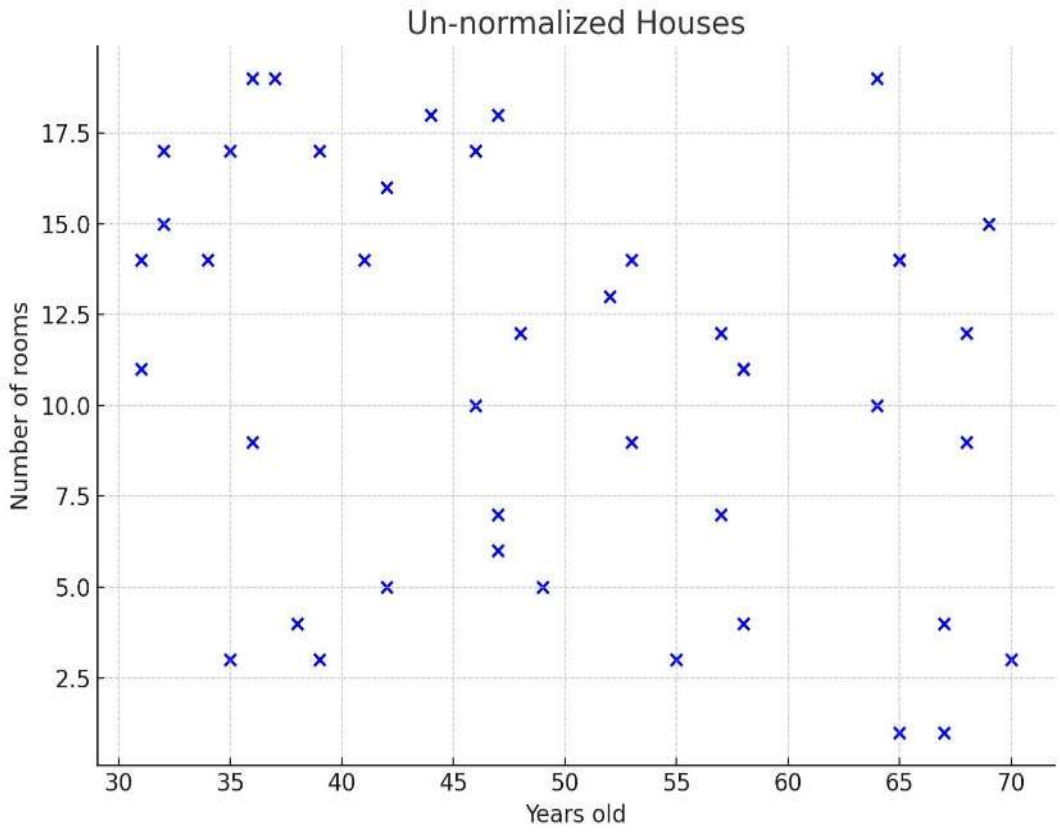

Figure 3.2 Data Set of House

By standardizing, we want to ensure that all data points are on the same scale and that each component is given the same weight. Below is an image that shows the same housing data that has been standardized using min-max standardization.

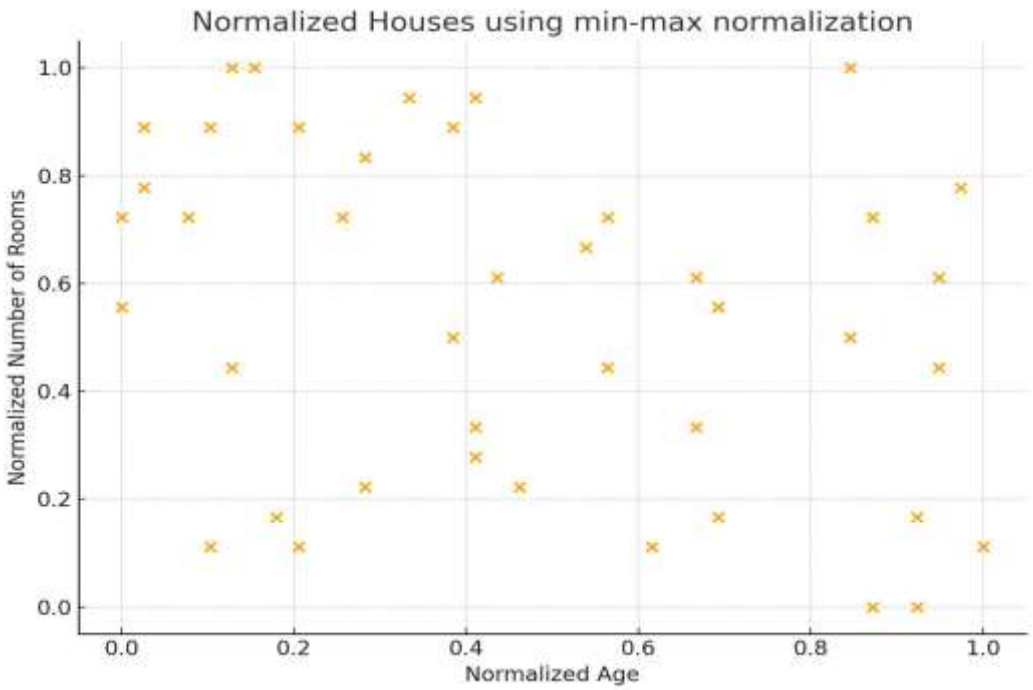

Figure 3.3 Normalized Houses using min-max normalization

- **MIN-MAX Normalization**

Among the many methods for standardizing data, min-max standardization is among the most well-known. The element's base estimate becomes zero, its maximum value becomes one, and all other values become decimal numbers between zero and one for each component.

For example, since it falls between 20 and 40, 30 would be reduced to about 0.5 if the element's base estimate was 20 and its maximum value was 40. The following is the equation

$$\frac{Value - \min}{\max - \min}$$

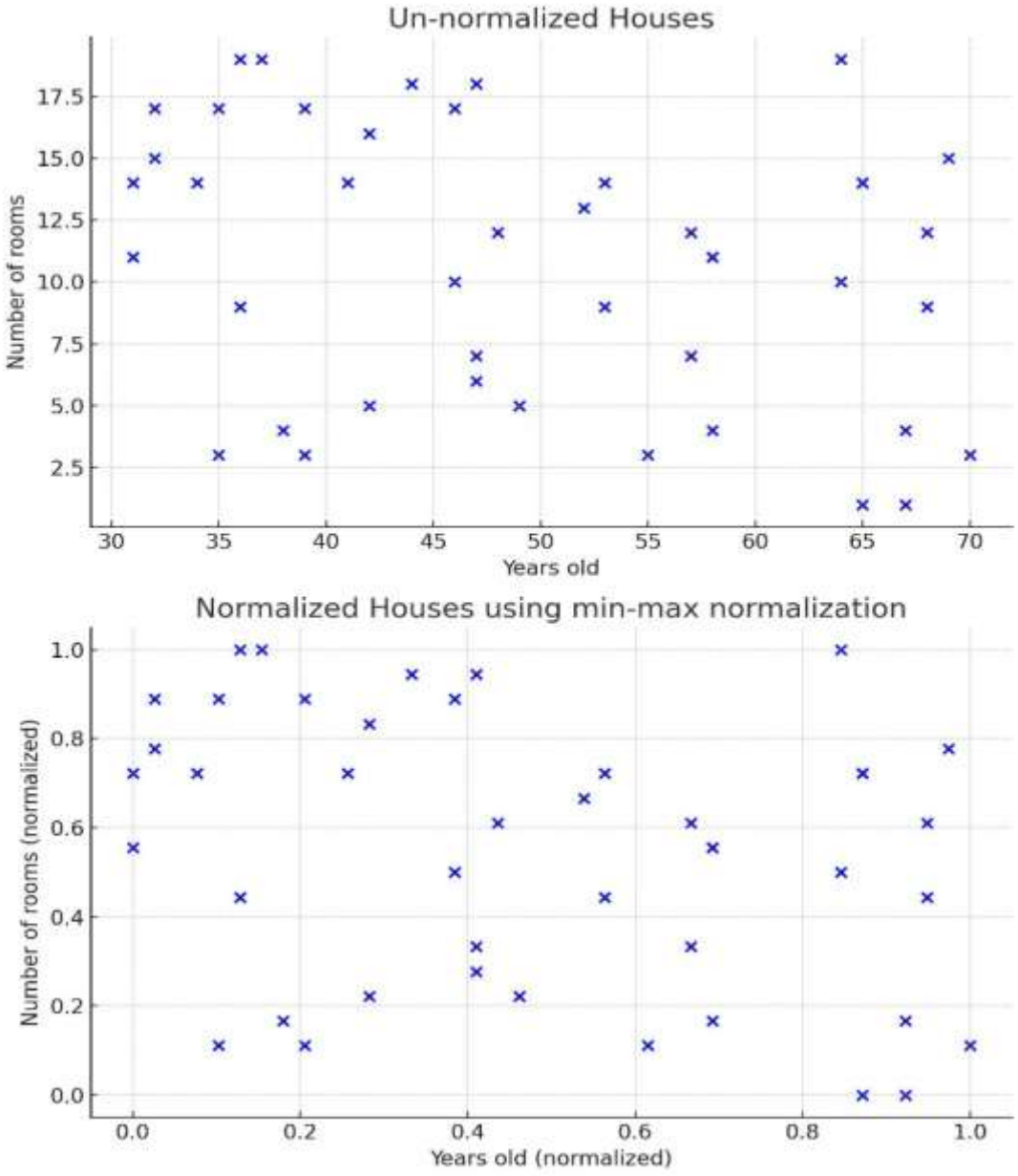

Figure 3.4 Normalized and Un-Normalized Houses

The x-axis remains troublesome, while normalizing resolved the y-axis squishing issue. Since the y-axis may vary by 1, while the x-axis can only differ by 0.4, the y-axis would clearly take centre stage in any comparison of these points. The purpose of normalization is to prevent numerical errors when computing and to prevent the dominance of aggregate properties at larger numerical scales over those at smaller ones.

A popular technique, Min-Max Normalization, is used in this study. By discovering, it converts a view v of the remarkable dataset to v' within the range of [new_min; new_max].,

$$V' = \left(\frac{V - V_{\min}}{V_{\max} - V_{\min}} \right) \times (\text{new}_{\max} - \text{new}_{\min}) + \text{new}_{\min}$$

New _min and new _max provide the range of values for the progress dataset in this case. In this test, we're connecting new_max= 1 and new_min= 0. The modified datasets are then used for the property decreasing system after normalization.

3.6 ATTRIBUTE REDUCTION BASED ON MODIFICATION IN THE FIREFLY ALGORITHM

Each firefly represents a candidate solution vector

$$x_i = [x_{i1}, x_{i2}, \ldots, x_{id}]$$

Brightness (objective value)

Brightness is proportional to the fitness of solution xi:

$$I_i = f(x_i)$$

For feature selection, the fitness combines accuracy and number of selected features:

$$F(x_i) = \alpha(1 - Acc(x_i)) + \beta \left(\frac{|S_i|}{|F|} \right)$$

where

- Acc(xi) = classification accuracy using subset SiS_iSi
- |Si| = number of selected features
- |F||= total features
- β = weighting constants (α+β=1\alpha+\beta=1α+β=1).

Attractiveness

Attractiveness of firefly j to firefly iii at distance rij is:

$$\beta(r_{ij}) = \beta_0 e^{-\gamma r_{ij}^2}$$

With

$$r_{ij} = \|x_i - x_j\|_2$$

Where β0 is initial attractiveness and γ is the light absorption coefficient.

$$x_i^{t+1} = x_i^t + \beta_0 e^{-\gamma r_{ij}^2}(x_j^t - x_i^t) + \alpha \varepsilon_i^t$$

where

- t = iteration index
- α = randomization parameter
- ε_i^t = random vector (e.g. uniform in $[-0.5, 0.5]$).

For **binary feature selection**, a sigmoid or threshold function is applied:

$$x_{ik}^{t+1} = \begin{cases} 1, & \text{if sigmoid}(x_{ik}^{t+1}) \geq \tau \\ 0, & \text{otherwise} \end{cases}$$

with $\tau \in (0, 1)$.

Algorithm workflow:

1. Initialize firefly population with feature subsets.
2. Evaluate brightness using RST dependency.
3. Move fireflies according to MFA rules.
4. Apply directional update to brightest firefly.
5. Repeat until convergence.
6. Return optimal reduct.

3.7 ATTRIBUTE REDUCTION BASED ON ROUGH SETS

Efficient attribute reduction is essential for building an accurate and computationally efficient heart disease prediction model. Medical datasets often contain redundant, irrelevant, or weakly correlated features, which may negatively influence classifier performance. To overcome this challenge, the proposed work integrates Rough Set Theory (RST) for evaluating feature significance and a Modified Firefly Algorithm (MFA) for searching the optimal attribute subset. The innovation of this method lies in the modification of the classical Firefly Algorithm: the brightest firefly (representing the best feature subset) is restricted to move only in directions that improve its fitness, avoiding random deterioration and significantly improving convergence and reliability. This is particularly important when working with high-dimensional clinical datasets where optimal feature selection directly impacts prediction accuracy.

3.6.1 Rough Set Theory for Attribute Reduction

Rough Set Theory (RST) is a mathematical tool used to reduce attributes without requiring any preliminary information, such as probability or membership values. Given a decision table:

$$v_i^{t+1} = wv_i^t + c_1 r_1 (pbest_i - x_i^t) + c_2 r_2 (gbest - x_i^t)$$

$$x_i^{t+1} = x_i^t + v_i^{t+1}$$

$T = (U, A \cup \{D\})$

Where:

- U = Universe of objects
- A = Set of conditional attributes
- D = Decision attribute

The dependency degree of D on C ⊆ A is defined as:

$$\gamma\_C(D) = |POS\_C(D)| / |U|$$

Where:

$$POS\_C(D) = \bigcup (\bar{C}(d)), \text{ for all } d \in D$$

A subset C′ is a reduct if:

$$\gamma\_\{C'\}(D) = \gamma\_C(D)$$

and removing any attribute from C′ decreases the dependency.
Thus, RST provides a mathematical basis for verifying whether selected attributes preserve classification quality.

3.4.2 Limitations of the Classical Firefly Algorithm

While the Firefly Algorithm is a powerful meta-heuristic optimization technique, its classical form suffers from two major drawbacks when applied to healthcare data:

1. Random Movement of the Brightest Firefly

In standard FA, even the best-performing firefly may move randomly. This can degrade its brightness (solution quality), slow convergence, and reduce the chances of reaching the optimal feature subset.

2. Premature Convergence

Medical datasets often contain correlated attributes. FA may get trapped in local optima and fail to explore promising regions effectively.

To address these limitations, the proposed method introduces a direction-controlled movement mechanism**,** which significantly enhances both convergence and solution stability.

3.8 PRINCIPAL COMPONENT ANALYSIS (PCA)

A principal component analysis (PCA) system's final output is a set of vectors in a space with dimensions shifted from one to another. Principal Component Analysis (PCA) is an element extraction approach that yields new highlights that are a linear mix of the underlying highlights; it uses this method to extract a reduced dimensional element subset. Its ultimate objective is to ensure that k is less than d by mapping all instances Convert the provided dataset from a d-dimensional space to a k-dimensional subspace.

The Principal Components (PC) are the k-new dimensions that are created, and each PC is coordinated to achieve the maximum change possible, excluding the difference that is already reflected in all of its initial segments. Thus, the main section accounts for the most difference, while each subsequent component accounts for a smaller estimate of volatility. The following is a way to refer to the Principal Components.

Let the dataset be represented as:

$X = [x1, x2, ..., xn]^T$

where each xi has d features.

Step 1: Standardize the Dataset

$x'ij = (xij - \mu j) / \sigma j$

Step 2: Compute the Covariance Matrix

$S = (1/(n-1)) * (X^T X)$

Step 3: Eigenvalues and Eigenvectors

$S * ei = \lambda i * ei$

Step 4: Construct the Transformation Matrix

$W = [e1, e2, ..., ek]$

Step 5: Transform the Dataset

$Z = XW$

3.5.3 Interpretation of Principal Components

PCi = ai1 x1 + ai2 x2 + … + aid xd

Here, PC_i Principal component I,

Xj – original feature ,,j ;

aj – numerical coefficient for Xj.

The following procedure may be used to calculate the primary components:

1. Take the data input and compute the S-covariance matrix.

2. Determine the eigenvalues and eigenvectors of S and arrange them in decreasing order according to the eigenvalues.

3. Using the preset number of components (eigenvectors), create the actual transition matrix.

4. Finally, to get a lower-dimensional representation, multiply the initial feature space by the obtained transition matrix.

3.9 MACHINE FOR SUPPORT VECTOR FUNCTIONS (RBF-SVM)

- **Basics of SVM**

A maximum separating hyperplane is built by mapping the input vector to a higher dimensional space. On each side of the data-splitting hyperplane, two parallel hyperplanes are drawn. A hyperplane that optimizes the distance between two parallel hyperplanes is termed the separating hyperplane. As a starting point, SVM uses

$$(x_1,y_1),(x_2,y_2),\ldots.(x_n,y_n)\ ,\ x_j \in R^n\ ,\ y_j \in \{+1,-1\}\ .$$

Here x_j is the input feature vector of j^{th} sample and y_j is the output index which is +1 or -1. SVM uses a hyperplane to split the positive and negative cases as

$$w..x+b=0, w\in R^n, b\in R$$

There is a gap between the positive and negative instances. Using a margin maximisation algorithm, SVM determines the optimal hyperplane.

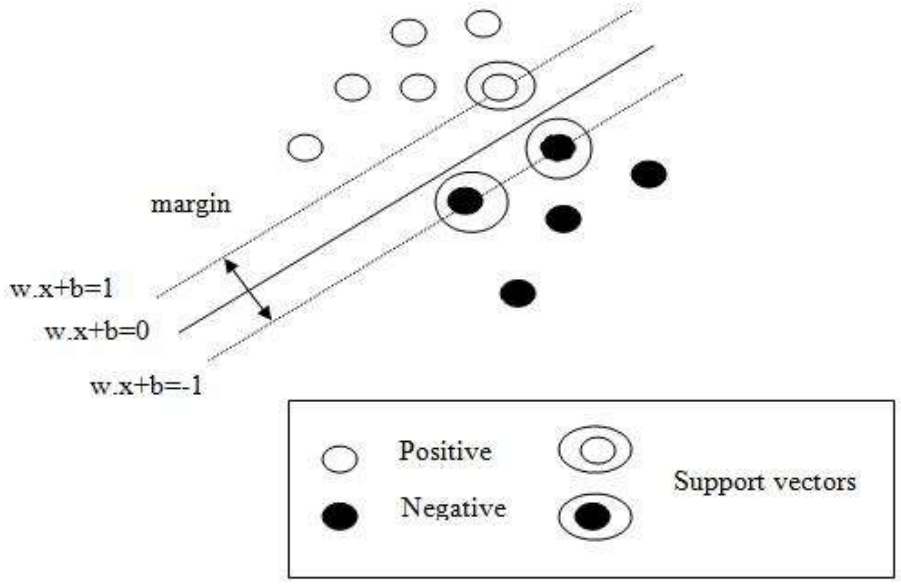

Figure 3.5 Support Vector Machines

$$w..x_j+b \geq 1-\xi_j$$

$$w..x_j+b \leq -1+\xi_j$$

To minimise the objective function stated in (3), we must take into account the restrictions of (2).

$$\frac{1}{2}\|w\|^2 + C\sum_{j=1}^{u}\xi_j$$

Both the margin size and the misclassification are indicated by the first and second components, respectively, of Eqn. (3). In this case, the cost of unmet limitations is represented by the changeable positive number C.

If the situation is linearly separable, the decision function f(x) = sgn(g(x)) is provided as

$$g(x) = \left(\sum_{i=1}^{l} \lambda_i y_i x_i . x + b \right)$$

The decision function in non-linear situations is provided by

$$g(x) = \left(\sum_{i=1}^{l} \lambda_i y_i K(x_i . x) + b \right)$$

where K K(xi .x) is a kernel function given by

$$K(x_i . x) = (x_i . x + 1)^d$$

- **Radial Basis Function (RBF)**
- **Radial Functions**

Starting with a basic issue and using great approaches, their reaction monotonically lowers (or grows), which is their characteristic highlight. In a direct model, the parameters (such as the inside of the separation scale and the precise condition of the spread capacity) remain constant. For scalar inputs, a common radial function is the Gaussian, which looks like:

$$h(x) = \exp\left(\frac{(x-c)^2}{r^2}\right)$$

The characteristics of this object are its centre (c) and radius (rr). With a centre of zero and a radius of one, a Gaussian RBF is shown in Figure 3.6 . As one moves out from

the centre, the radius of a Gaussian RBF monotonically shrinks. With scalar input, on the other hand, a multiquadric RBF is

$$h(x) = \frac{\sqrt{r^2 + (x-c)^2}}{r}$$

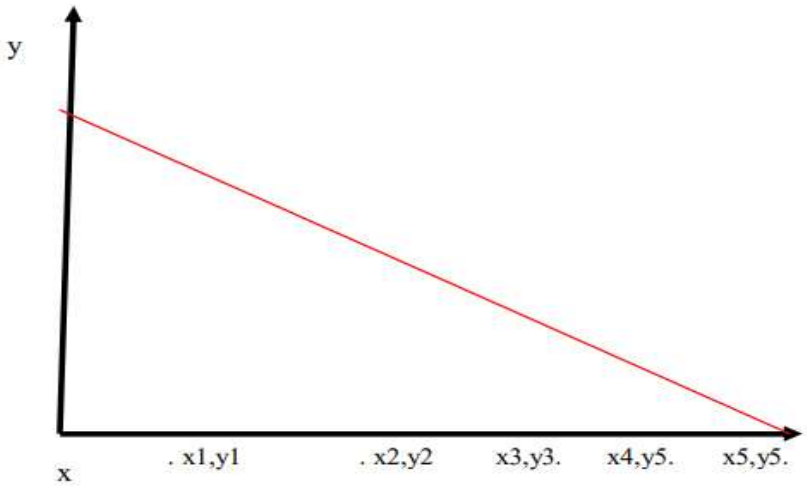

Figure 3.6 Fitting a straight line to a bunch of points is a kind of parametric regression where the form of the model

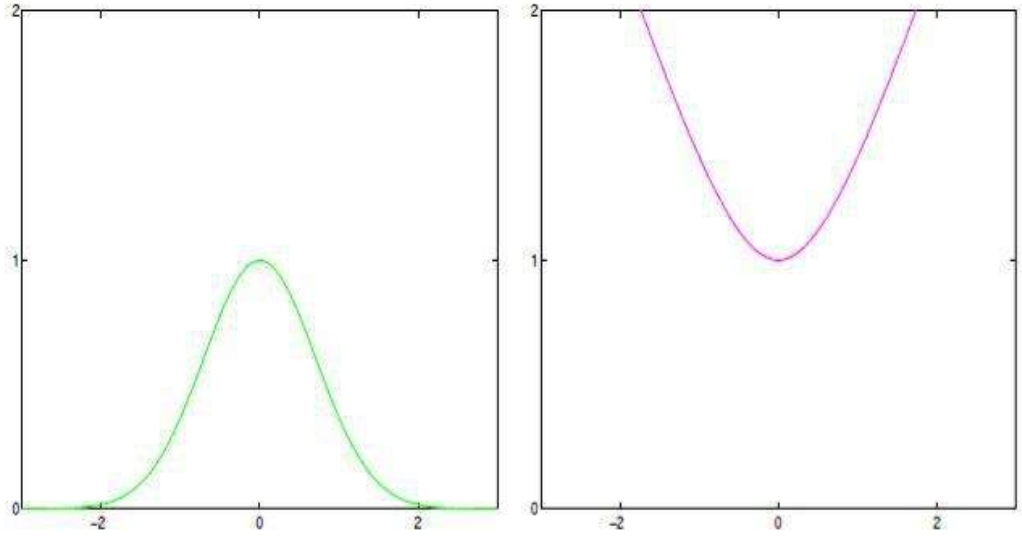

Figure 3.7 Gaussian (left) and Multiquadric RBFs

- **RBF Neuron Activation Function**

In RBF networks, each neurone analyses data based on a percentage of how similar it is to its model vector, which is drawn from the training set. Data vectors that are more and more similar to the model provide results closer to 1. Although there are a number of possible judgements on comparability capabilities, the most famous one is based on the Gaussian. In the case of a one-dimensional information, the following holds for a Gaussian.

The information is represented by x, A standard deviation is mu times the mean is sigma. This results in the naturally occurring chime bend seen below, which is centred on the mean, mu (where 5 is the mean and 1 is sigma) in the figure below.

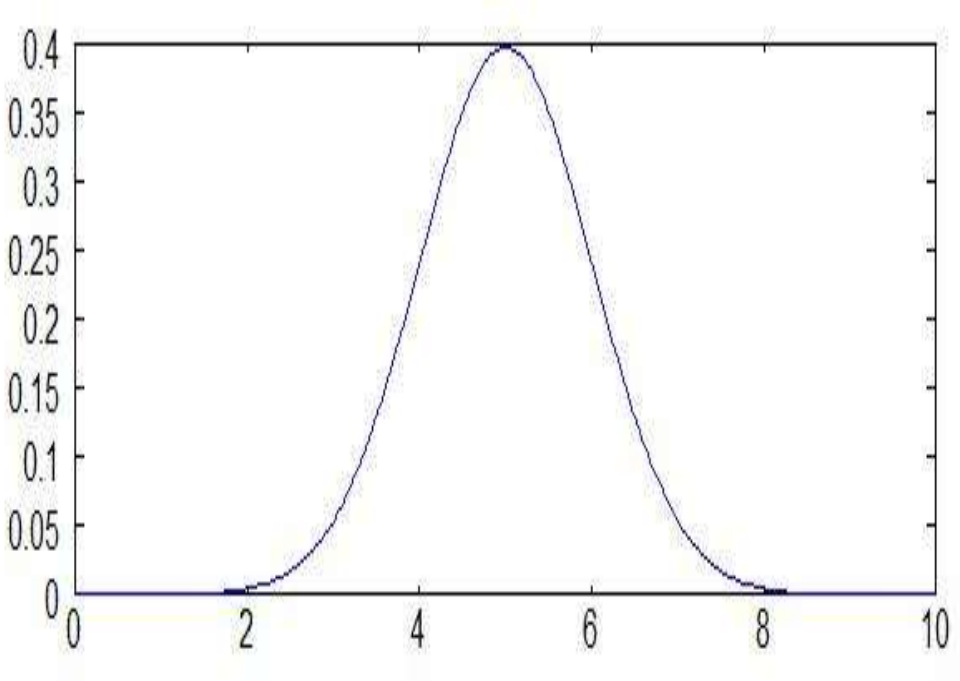

Figure 3.8 Familiar Bell curve

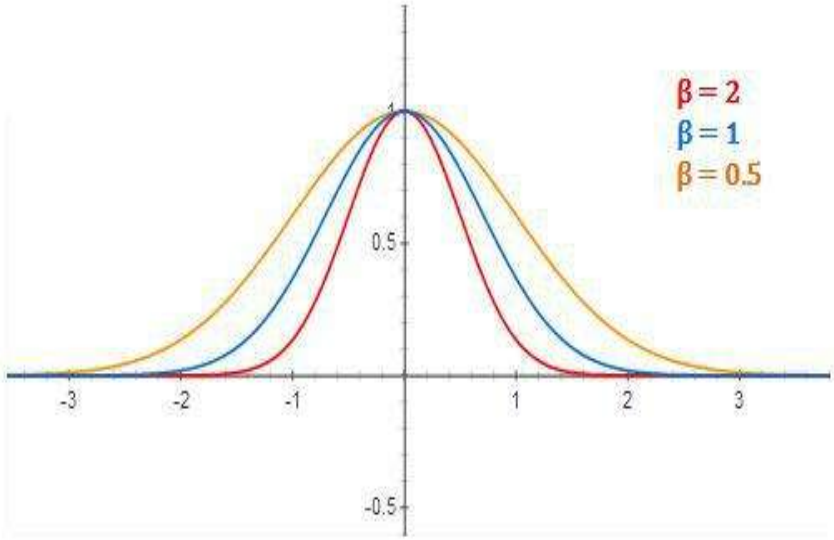

Figure 3.9: RBF Neuron activation for different values of beta

The Gaussian component, which is based on RBF, transforms the space of the lower dimensions into an unfathomably high-dimensional space. Unidentifiable highlights that are projected into three-dimensional space always end up being vividly visible.

$$KF(x'', x_i'') = (-\gamma \left\| x'''', x_i'' \right\|)^2 : \forall$$

The γ is responsible for adjusting the Gaussian ringer mold's width. The smaller the estimate, the wider the curve, and vice versa. When the RBF component is combined with SVM, the final result is a step closer to becoming

$$f(x'' = \sum_{i=1}^{m} \alpha_i e^{(-\gamma \left| x'' \cdot x_i \right|)} + b$$

The RBF-based support vector machine (SVM) has two categories: normal subjects (NS) and heart patients (HP).

3.10 IMPROVING HEART DISEASE PREDICTION WITH TRANSDUCTIVE SUPPORT VECTOR MACHINE CLASSIFIER AND OPTIMAL USE OF WHOLE-SAMPLE OPTIMIZATION

Predicting cardiac problems using a Support vector machines and the radial basis function was the focus of the prior chapter. Having said that, the categorization result it produces is far from good. This chapter suggests a PSO-RS using TSVM — a combination of PSO and Rough Sets — as a solution to this challenge.

In this study, Zero-Score (Z-Score) is used for data normalization in order to decrease data redundancy and increase data integrity. To minimises computing cost and boost prediction system performance, the ideal subset of attributes is selected using the PSO algorithm and an attribute reduction approach based on Rough Sets (RS). Last but not least, Predicting cardiac sickness is done using the RBF-TSVM classifier.

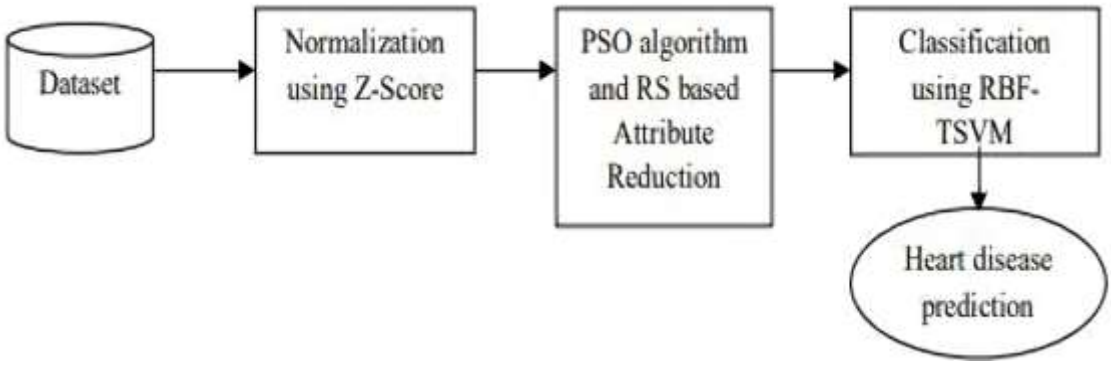

Figure 3.10 Block diagram of the proposed methodology

There are three main phases to the comprehensive design of a system for diagnosing heart disease: Classification, feature extraction, attribute reduction, and normalization.

- **Normalization Using Z-Score Normalization**

To avoid having data that are too close to one another in terms of distance measure, all of the input and output data were normalized before the testing and training operations. As far as Z-Score standardization is concerned, it is generally useful. This arrangement of scores is used to isolate each score by its standard deviation in order to standardize several scores using the normal deviation. In this specific case, before dividing by the standard deviation, we usually take the average score and deduct it from every single

score. The acronym Z-scores describes this uniformity [15]. According to official statistics, Z-scores are transformed into Yn for large N scores with means of M and standard deviations of S.

$$Z_n = \frac{Y_n - M}{S}$$

The standard deviation is 1 and the mean is 0 for many Z-scores which may be shown to be rather likely. Thus, Z-scores provide a unit-free metric that may be used to compare estimated perceptions with different units. The quality reduction approach is used to modified datasets after standardization.

Typically, the disagreeable set and data hypothesis is used in conjunction with the characteristics decline hypothesis. A reduction in qualities indicates a weakening of the knowledge base's repeating properties without crossing the line into characterization.

- **Attribute dependency**

The information table describes the choice property D's dependence on the condition characteristic C as:

$$\gamma(C,D) = \frac{|Posc(D)|}{|U|}$$

(4.2)

A positive domain's element count is given by |Posc(D)|.

Procedure characteristic decrease

1) Calculate the equality set of the condition quality set (C-{Ci});

2) Calculate the equality set of the choice quality set D;

3) Calculate the positive space Pos(C-{Ci} D);

4) Calculate the reliance γ(C-{Ci} D);

5) Calculate the significance of Ci: SGF (C-{Ci} D);

6) Reducing the attributes of which the importance is 0.

Data frameworks make extensive use of databases. The articles are categorized into identical sets according to the choice characteristic in the database. The trust is used to differentiate each choice class based on condition features. Finally, decision guidelines are generated for each class. There are a handful of data attributes that don't matter much for the learning task, but I have faith in finding a foundational set of correlative characteristics that can characterised the choice quality with the full set of conditions, and the rules I've been able to construct from this foundational set are getting easier and better.

3.11 A ROUGH SET ALGORITHM FOR REDUCING ATTRIBUTES USING PSO

- **Particle Swarm Optimization (PSO)**

A computational intelligence optimization method, the PSO methodology primarily draws inspiration from the habits of swarming or flocking animals like fish and birds.

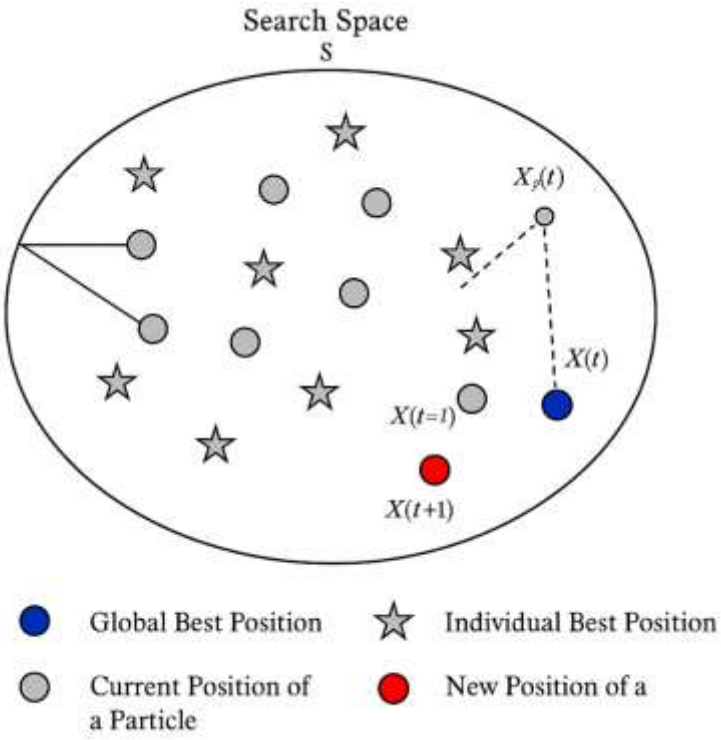

Figure 3.11 Functions of PSO

X_i denotes where the particle is located Yi stands for the particle's speed, and LS for the local memory space. GS stands for the space for all memories on Earth.

Equations (5.1) and (2) provide the particle's updated location, which may be used to calculate its velocity.

$$Y_i = Y_i + \beta_1 * rand * (L_{bSi} - X_i) + \beta_2 * rand * (G_{bSi} - X_i)$$

(4.4)

$$X_i = X_i + Y_i$$

where β_1 and β_2 be the constant that computes

$$Y_i = \gamma . Y_i + \beta_1 * rand * (L_{bSi} - X_i) + \beta_2 * rand * (G_{bSi} - X_i) \quad (4.6)$$

$$Y_i = Z * (\gamma * Y_i + \beta_1 * rand * ((L_{bSi} - X_i) + \beta_2 * rand * (G_{bSi} - X_i)) \quad (4.7)$$

Where $Z = \dfrac{2}{|2 - \beta - \sqrt{\beta^2 - 4\beta}|}$

and

$$\beta = \beta_1 + \beta_2$$

PSO Algorithm

---

1. Let S be the search space
2. Let X_i be the set that denotes the position of particle in S
3. Visualize Vi as the collection of nodes' velocities.
4. Let t_{FF} be the time delay between two successive fitness function assessments of a particle
5. Particles are initialized in S at position x_i\
6. Every particle computes the Fitness Function F()
7. After t_{FF}, F() of each particle is compared with Lbest
8. If ($F_{xi1}(t_{FF}1)$ >Lbest (x_{i1})) then
9. Lbest (x_{i1}) = $F_{xi1}(t_{FF}1)$

10. Else

11. Lbest (x_{i1}) is not modified

12. End if

13. After t_{FF}, F() of each particle is compared with Gbest

14. If ($F_{xi1}(t_{FF}1)$ >Gbest(x_{i1})) then

15. Gbest(x_{i1}) = $F_{xi1}(t_{FF}1)$

16. Else

17. Gbest(x_{i1}) is not modified

18. End if

19. Conditions (4.6) are satisfied, and the molecule's speed is restored (20).

20. The molecule's location is restored in accordance with condition (4.7).

21. Transfer the molecule to an other location

22. Keep going until the tree is full, then repeat steps 6–20.

- **Rough Set Algorithm**

By using this technique, the reduced

$$v_{ij}(t+1) = \omega * v_{ij}(t) + c_1 R_1 * \left(p_{best}(t) - X_{ij}(t)\right) + c_2 R_2 * \left(p_{best}(t) - X_{ij}(t)\right)$$

$$X_{ij}(t) = \begin{cases} 1, & \rho < s(v_{ij}(t), \\ 0, & otherwise \end{cases}$$

(4.8)

A weight factor value between 0.4 and 0.9 will improve the calculation's performance [20]. The following is the formula for the fluffy capacity s(v_ij (t)) that is often used in neural systems; it may be any quantity between 0 and 1:

$$s(v_{ij}(t) = \frac{1}{1+e^{-v_{ij}(t)}} \quad (4.9)$$

p_{best} and g_{best} represent the global ideal solution and the individual extreme solution, respectively, as shown in Equations (9) and (10):

$$p_{best} = \max(p_{best}, fitness(i)) \quad (4.10)$$

$$g_{best} - \max(p_{best}, g_{best}) \quad (4.11)$$

Wellbeing(I) is the wellbeing of molecule I in Eq. (4.10). The health metric serves as a primary guidepost for the swarm of molecules to follow as they strive for perfection. Since our decrease calculation is based on positive district decrease computation [21], we can easily acquire a base decrease with diverse consequences by changing the wellness capacity. To verify, check whether the relative positive location in the necessary conditions of Eq. (4.12) is true.

$$fitness(i) = \begin{cases} |i|, if \left(pos'_{|i|}(D)\right) = U'_{POS} \\ |c|, if \left(pos'_{|i|}(D)\right) \neq U'_{POS} \end{cases} \quad (4.12)$$

3.12 CLASSIFICATION USING RBF-TSVM

Heart disease prediction makes use of RBF-based TSVM support classification. A basic overview of the RBF approach was covered in the preceding chapter.

It is possible to combine the proposed broadcast data of unlabelled examples with well-prepared tests since TSVM computations make use of the potential of transductive adaptation effectively. Calculating using TSVM yields better grouping accuracy when compared to the standard assist vector machine method. Still, there are a number of issues with TSVM calculations. One of them is that N worth is sometimes difficult to obtain a reasonable estimate of, Also, for TSVM calculations, it's deceptive to show the number of positive name tests in the unlabeled samples. [22].

The ratio of positive tests to all unlabeled cases is evaluated by comparing the ratio of positive tests to all named tests; this ratio is then used as an estimate of N in TSVM calculations. Anyway, it's difficult to get a more precise estimate of N using this method when the number of exams with marks is small. If the number of tests with positive marks differs considerably from the pre-set estimate of N, then the TSVM computation will fail presentation becomes very weak, and the calculation's grouping accuracy cannot be effectively guaranteed.

$$(x_1, y_1), \ldots . (x_n, y_n), x_i \in R^m, y_i \in \{\ 1, \mid 1\}$$

with an additional set of unlabeled samples from the corresponding sharing,

$$x_1^*, x_2^*, x_3^*, \ldots \ldots, x_k^*$$

$$(y_1^*, \ldots \ldots y_k^*, w, b, \xi_1, \ldots \ldots, \xi_n, \xi_1^*, \ldots \ldots \xi_k^*)$$

$$\frac{1}{2}||w||^2 + C\sum_{i=1}^{n}\xi_i + C^*\sum_{j=1}^{n}\xi_j^*$$

Subject to:

$$\forall_{i=1}^{n}: y_i t[w.x_i + b] \geq 1 - \xi_i$$

$$\forall_{j=1}^{k}: y_j^*[w.x_j^* + b] \geq 1 - \xi_j^*$$

$$\forall_{i=1}^{n}: \xi_i \geq 0$$

$$\forall_{j=1}^{k}: \xi_j^* \geq 0$$

Training in TSVM

Classify the test examples using <~w;b>. The num+test examples with the highest value of $\vec{w} * \vec{x}_j^* + b$ are assigned to class $+(\vec{y}_j^* := 1);$;

Students are given the remaining test cases to complete in class.

$$-+(\vec{y}_j^* := -1);$$

$$C_-^* := 10 - 5;//$$:

.

some small number

$C_-^* := 10 - 5 * \frac{num+}{k-num+}$;

While $((C_-^* < C^*) || (C_-^* < C^*))$ {

//Loop1

$(\vec{w}, b, \vec{\xi}, \vec{\xi}^*) :=$
$solve\_sum\_qp([(\vec{x_1}, y1), \ldots, (\vec{x_n}, yn)], [(\vec{x}_1^*, \vec{y}_1^*), (\vec{x}_k^*, \vec{y}_k^*)], C, C_-^*, C_+^*)$
;

1) While $(\exists m, 1 : (y_m^* * y_1^* < 0) \& (\xi_m^* > 0) \& (\xi_1^* > 0) \& (\xi_m^* * \xi_1^* > 2))$

{

Loop2

$y_m^* := -y_m^*$; //take a positive and a negative test

$y_1^* := -y_1^*$; // example, switch their labels, and retrain

$(\vec{w}, b, \vec{\xi}, \vec{\xi}^*) :=$
$solve\_sum\_qp([(\vec{x_1}, y1), \ldots, (\vec{x_n}, yn)], [(\vec{x}_1^*, \vec{y}_1^*), (\vec{x}_k^*, \vec{y}_k^*)], C, C_-^*, C_+^*)$
;

}

$C_-^* := \min(C_-^* * 2, C^*)$;

$C_+^* := \min(C_+^* * 2, C^*)$;

}

return $(\vec{y}_1^* - \vec{y}_k^*)$;

The following are the main components of a TSVM training algorithm:

TSVM training Algorithm

Input: -training examples $(\overrightarrow{x\ 1}, y1), \ldots., (\overrightarrow{x\ n}, yn)$

-test examples $\vec{x}_1^*, \ldots \ldots \vec{x}_k^*$

Parameters: -C,C*: parameters from OP(2)

-num+: class's required amount of practice exams+

Output: -anticipated labels for the sample data

$$\vec{y}_1^* - \vec{y}_k^*(\overrightarrow{w}, b, \vec{\xi}, \quad) := solve\_sum\_qp([(\overrightarrow{x\ 1}, y1), \ldots., (\overrightarrow{x\ n}, yn)],$$

Stage 1: Find C and C*, finish an underlying learning using inductive picking up using all marked cases, and construct a one-of-a-kind classifier. In the unlabelled models, identify a positive-named model by assigning it a predicted value N.

Stage 2: Evaluate each unlabelled segment using the first classifier to determine its option capacity charges. All but one of the models that include Mark N's actual judgement abilities are considered detrimental by him. Establish a short-term achieve factor C_tmp^*.

Stage 3: In most cases, you should retrain the support vector machine. Alter the names of a few different named unlabelled representations according to a specified guideline in order to estimate the goal capacity drop as much as feasible for the recently produced classifier. This process continues until no two models that satisfy the exchange condition have been constructed.

Stage 4: Return to Step 3 after slightly increasing the estimate of C_tmp^*.

Whenever C_tmp^* C∗, the calculation is finished and the outcome is yield.

It is confirmed that the target capacity would decrease after the exchange using the marker swapping approach in Step 3. Step 4's iteratively increasing temporary accomplish factor calculates the influence of the unlabelled models on the goal task gradually in an effort to look for a convenient error supervise. Due to the fact that C* provided in Step 1 is a finite quantity, the computation might conclude after limited cycles.

The radial basis function (RBF) based component, often known as the Gaussian piece, transforms the space of lower dimensions into an infinite dimensional space. It is common for straightly non-divisible highlights to become directly detachable after being transferred into higher dimensional space.

Predicting the occurrence of cardiac illnesses is one use of the RBF-based TSVM for classification purposes.

3.13 METHODOLOGICAL COMPARISON BETWEEN EXISTING AND PROPOSED MODELS

Table 3.4 : Methodological Comparison Between Existing and Proposed Models

Criteria	Existing Model (IT2FLS)	Proposed Models (MFA–RBF-SVM, PSO–RBF-TSVM, OCSO–RBF-SVM)
Feature Selection	No systematic feature selection; uses all features	Uses MFA and PSO-RS for optimal feature subset selection
Handling Nonlinearity	Limited ability to model nonlinear patterns	RBF-SVM and TSVM effectively capture nonlinear decision boundaries
Parameter Optimization	Parameters manually chosen; no optimization framework	MFA and OCSO provide automatic hyperparameter tuning
Adaptability to Noisy Data	Sensitive to noise due to fixed fuzzy rules	Robust optimization reduces noise impact and improves stability
Flexibility	Rule-based, less adaptable to complex datasets	Hybrid models dynamically adapt to feature interactions

Computational Efficiency	Moderate; rule evaluation only	Improved efficiency with reduced feature subsets and optimized parameters
Generalization Ability	Weak generalization due to static rule base	Strong generalization through optimized classifiers and feature subsets
Scalability	Limited scalability to increasing data size	Models scale effectively with larger datasets due to optimization strategies
Learning Capability	No learning mechanism; rule-driven	Machine-learning + optimization allows continuous improvement
Suitability for Clinical Prediction	Basic interpretability but limited accuracy	High adaptability, strong predictive power, and better clinical reliability

The comparison clearly shows that IT2FLS lacks structured feature selection, parameter optimization, and the ability to model nonlinear relationships. In contrast, the proposed hybrid models leverage optimization algorithms and advanced classifiers to deliver greater adaptability, improved generalization, and stronger methodological foundations for heart-disease prediction.

3.14 EXPERIMENTAL RESULTS AND ANALYSIS

The performance of the proposed hybrid approaches—MFA–RBF-SVM **and** PSO–RBF-TSVM—is evaluated using standard classification metrics: Accuracy**,** Sensitivity**,** Specificity**,** False Positive Rate (FPR) **and** False Negative Rate (FNR)**.** The experiments are performed using different dataset sizes ranging from 1000 to 5000 records**,** derived from the Cleveland Heart Disease Dataset (CHDD)**.**The results are compared with the existing Interval Type-2 Fuzzy Logic System (IT2FLS), which serves as the baseline model.

FALSE POSITIVE RATE (FPR)

The False Positive Rate represents the proportion of healthy individuals who are incorrectly classified by the model as having heart disease. A high FPR leads to unnecessary clinical tests, patient anxiety, and increased healthcare costs.

$$\text{FPR} = \frac{FP}{FP + TN}$$

Where:

- **FP (False Positives):** Healthy patients misclassified as diseased
- **TN (True Negatives):** Healthy patients correctly classified

FALSE NEGATIVE RATE (FNR)

The False Negative Rate indicates the proportion of actual heart disease cases that the model fails to detect. A high FNR is dangerous because it may delay diagnosis and treatment, potentially resulting in severe medical complications.

$$\text{FNR} = \frac{FN}{FN + TP}$$

Where:

- **FN (False Negatives):** Heart disease patients misclassified as healthy
- **TP (True Positives):** Heart disease patients correctly classified

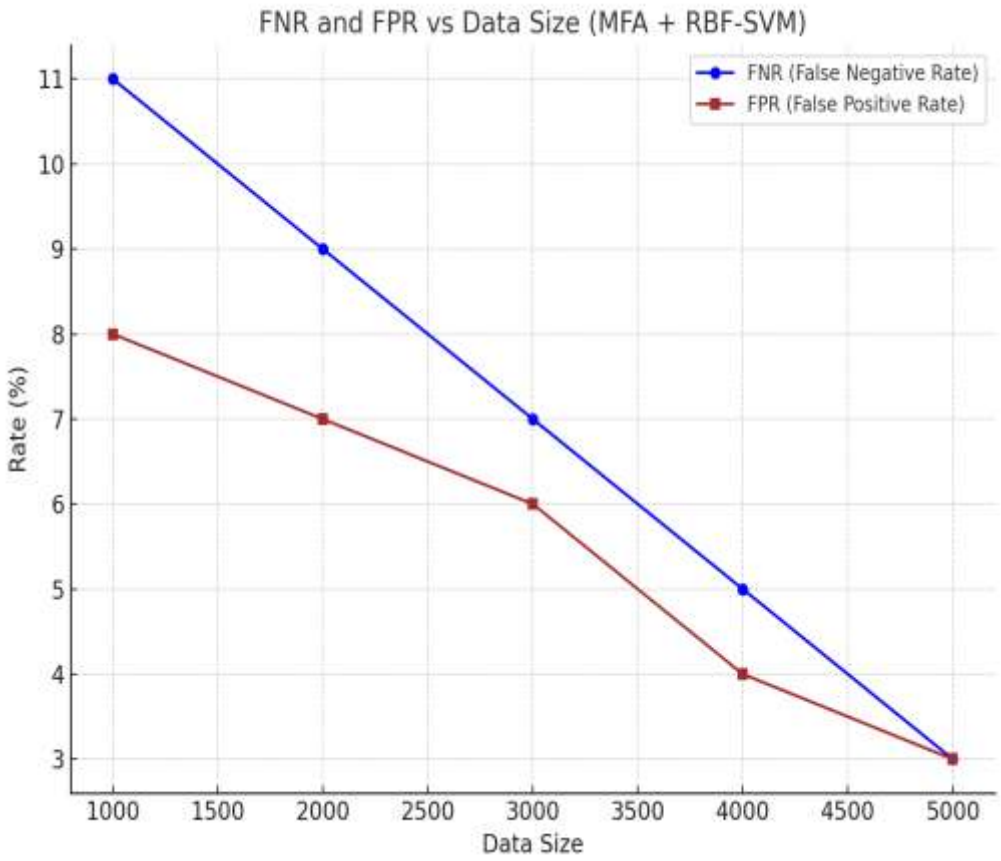

Figure 3.12 : Graph FPR and FNR

This graph illustrates the relationship between data size and the error rates—False Negative Rate (FNR) and False Positive Rate (FPR)—for the MFA + RBF-SVM model. As the amount of training data increases from 1000 to 5000 records, both FNR and FPR consistently decrease. The FNR, represented by the blue line, drops from 11% to 3%, indicating that the model is making fewer mistakes in missing actual positive cases. Similarly, the FPR, shown in brown, decreases from 8% to 3%, showing a reduction in incorrect positive predictions. This trend demonstrates that increasing the data size leads to better model performance, with fewer classification errors and improved accuracy in both identifying true positives and true negatives.

(i) ACCURACY – MFA–RBF-SVM VS IT2FLS

The measurement precision allows for the proper segmentation of tumour components in pictures based on their weighted ratio. It is represented in this way,

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \times 100$$

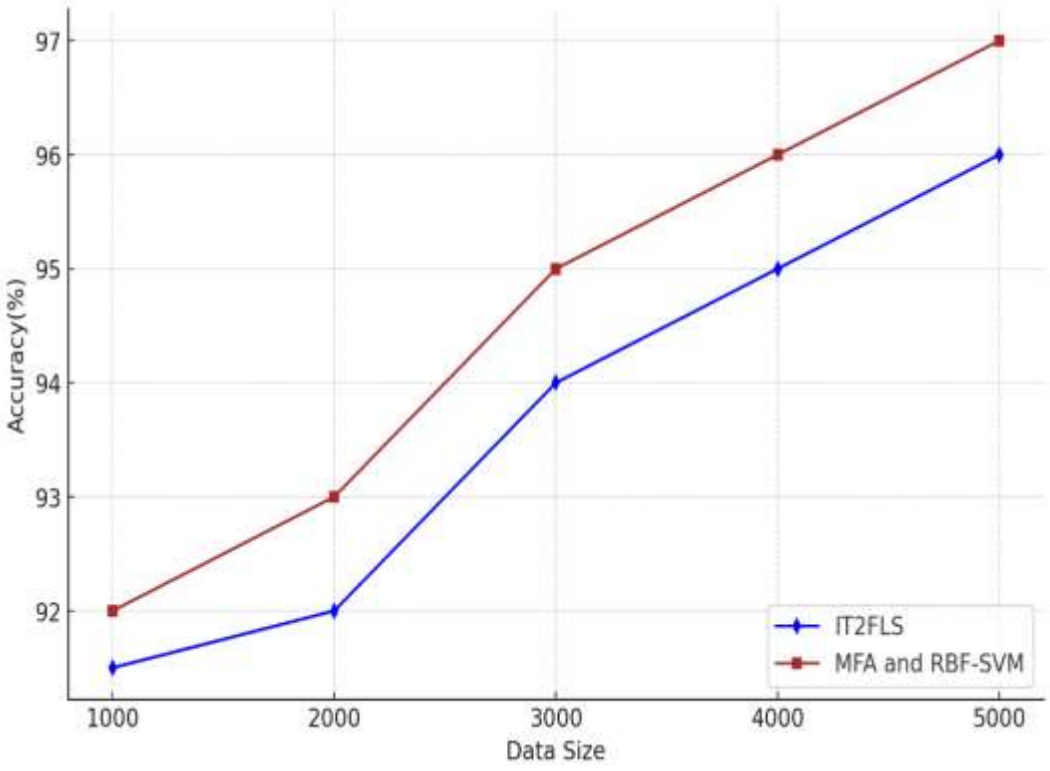

Figure 3.13: Accuracy Comparison

Figure 3.12 compares the accuracy of the proposed **MFA–RBF-SVM** method with the existing **IT2FLS** method across different dataset sizes. The x-axis represents the dataset size (number of records), and the y-axis represents accuracy (%).Using MFA for attribute reduction before classification allows the proposed system to achieve higher accuracy at all dataset sizes. The MFA–RBF-SVM model consistently outperforms IT2FLS, demonstrating the effectiveness of the optimisation-based feature selection.

(ii) SENSITIVITY – MFA–RBF-SVM VS IT2FLS

A high level of affectability is indicative of a large number of accurately recognized positive aspects. In order to perceive favourable results, it identifies with the test's boundary.

$$\text{Sensitivity} = \frac{TP}{TP + FN} \times 100$$

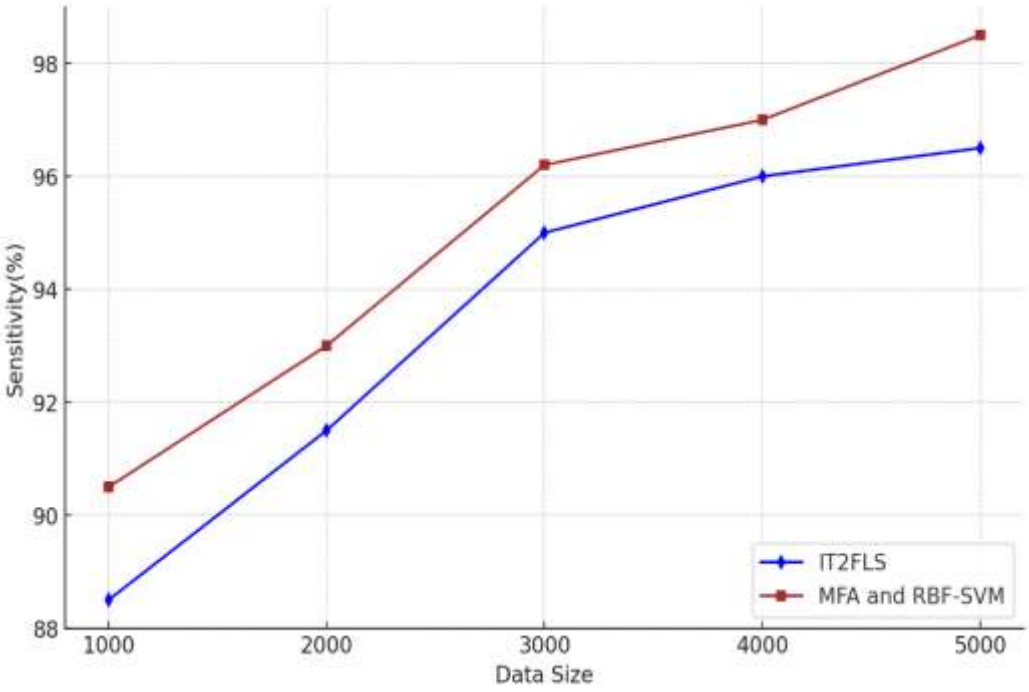

Figure 3.14: Sensitivity Comparison

Figure 3.13 shows the sensitivity of the IT2FLS method and the proposed MFA–RBF-SVM method for different dataset sizes. The y-axis represents sensitivity, and the x-axis shows dataset size.

The proposed work uses **min–max normalization** and **MFA-based attribute reduction**, followed by RBF-SVM classification. This combination improves the detection of true heart disease cases and hence increases sensitivity. Across all dataset sizes, MFA–RBF-SVM achieves higher sensitivity than IT2FLS, meaning fewer heart disease cases are missed.

(iii) SPECIFICITY – MFA–RBF-SVM VS IT2FLS

An indicator of specificity is the fraction of false negatives that are properly classified. The test's capacity to identify unfavourable outcomes is connected to it.

$$\text{Specificity} = \frac{TN}{TN + FP} \times 100$$

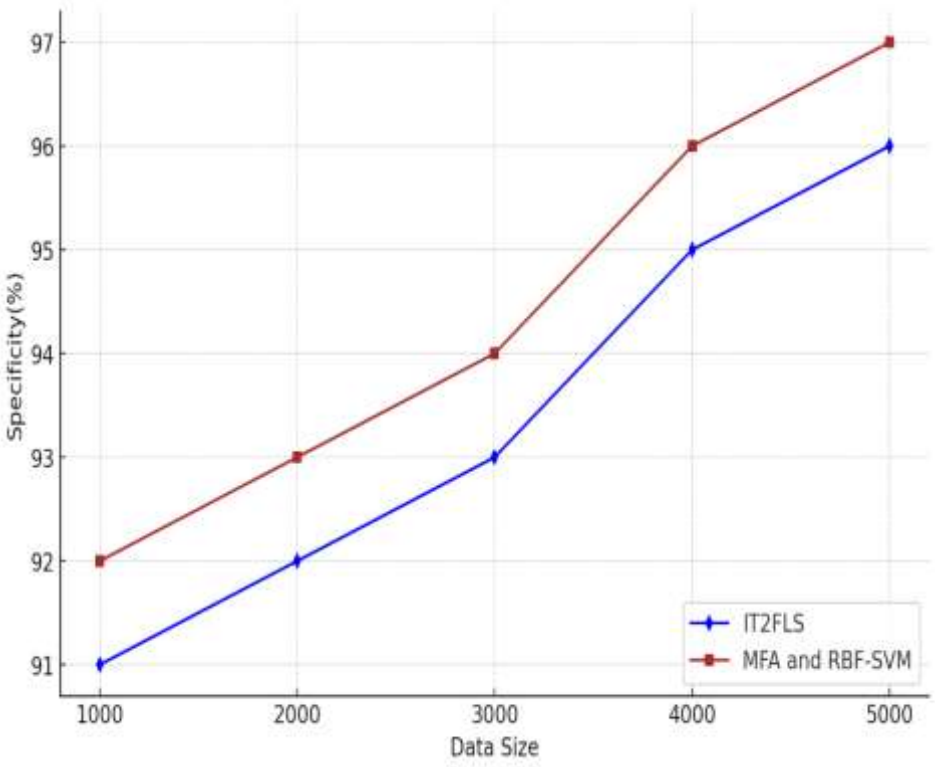

Figure 3.15: Specificity Comparison

Figure 3.14 compares the specificity of IT2FLS and MFA–RBF-SVM. The x-axis shows dataset size, and the y-axis shows specificity.

The results indicate that MFA–RBF-SVM yields higher specificity than IT2FLS for all dataset sizes, implying that the proposed model produces fewer false alarms (healthy subjects wrongly predicted as diseased) and is more reliable in identifying non-diseased cases.

Table 3.5 Results of MFA-RBF-SVM and IT2FLS

Data Size (Bytes)	Accuracy		Sensitivity		Specificity	
	IT2FLS	**MFA and RBF-SVM**	**IT2FLS**	**MFA and RBF-SVM**	**IT2FLS**	**MFA and RBF-SVM**
1000	91.5	92	87	89	91	92
2000	92	93	89	91	92	93
3000	94	95	91	93	93	94
4000	95	96	94	95	95	96
5000	96	97	95	97	96	97

The table compares the performance of IT2FLS and MFA with RBF-SVM using accuracy, sensitivity, and specificity across data sizes ranging from 1000 to 5000 bytes. Although both models improve as the data size increases, MFA with RBF-SVM consistently achieves higher performance at every level. IT2FLS shows accuracy improving from 91.5% to 96%, sensitivity from 87% to 95%, and specificity from 91% to 96%. In comparison, MFA with RBF-SVM increases accuracy from 92% to 97%, sensitivity from 89% to 97%, and specificity from 92% to 97%. Overall, the results indicate that while both models benefit from larger datasets, MFA with RBF-SVM provides superior classification performance.

Table 3.6 The MFA-RBF-SVM Outperforms IT2FLS in Percentage Terms

Data Size	Accuracy	Sensitivity	Specificity
1000	0.54	2.24	1.08
2000	1.07	2.19	1.07
3000	1.05	2.15	1.06
4000	1.04	1.05	1.04
5000	1.03	2.06	1.03

Cleveland Heart Disease Dataset (CHDD), which can be found on the UCI Repository [14], is the source of the datasets. Here are the thirteen characteristics that are taken into account: details about the patient's age, gender, the nature of their chest discomfort, blood pressure at rest, cholesterol levels, glucose levels after fasting, electrocardiogram readings at rest, angle of the heart, number of major veins obscured by fluoroscopy, and the extent to which exercise-induced ST depression differs from resting conditions. A range of dataset sizes (from 1000 to 5000 records) is considered for performance evaluation of the proposed methods. Here, we evaluate the sensitivity, specificity, and accuracy of the suggested PSO and RBF-TSVM method with that of the current system IT2FLS [23], as well as with modified FA and RBF-SVM

(i) Accuracy-PSO–RBF-TSVM vs IT2FLS and MFA–RBF-SVM

Accurate measurement allows for the proper segmentation of the weighted proportion of tumour sections in pictures. This is shown as,

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \times 100$$

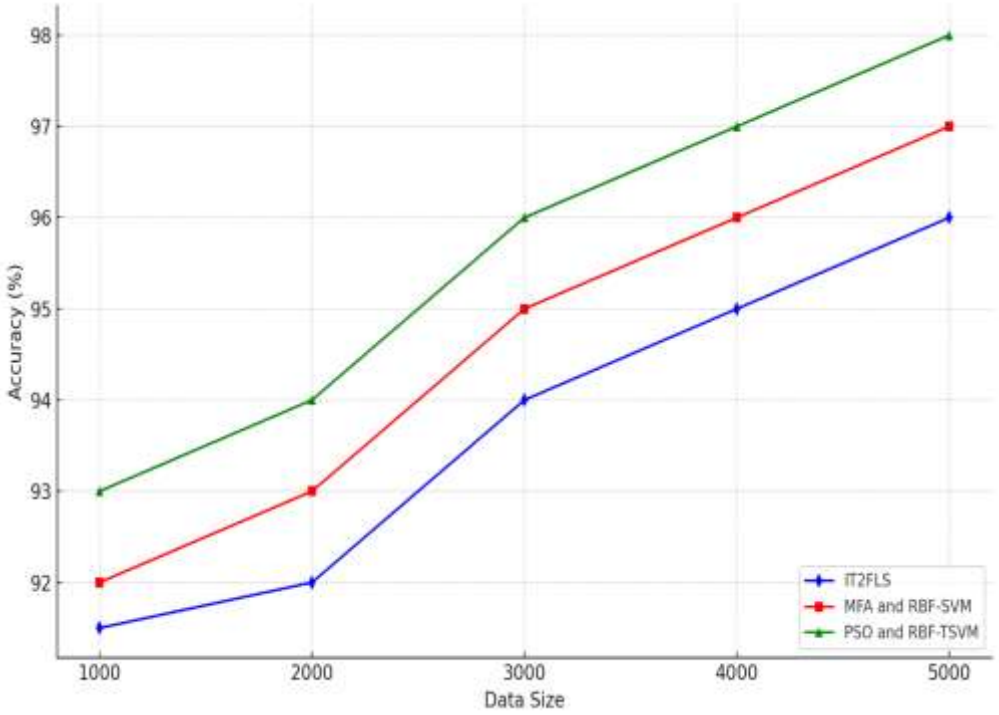

Figure 3.16. Results for Accuracy

Comparing the accuracy of the current IT2FLS, MFA, and RBF-SVM based classification methods with the suggested PSO-RBF-TSVM based methodology is shown in Figure 3.15. On the one hand, we have the dataset size (X-axis) and accuracy (Y-axis). The suggested approach employs PSO for attribute reduction to get high accuracy. When compared to the current technique, the PSO and based RBF-TSVM classification algorithms demonstrated very high accuracy across all dataset sizes

(ii) Sensitivity – PSO–RBF-TSVM vs Baseline Models

Affectability is defined as the degree to which pleasant emotions are adequately felt. In order to perceive favorable results, it identifies with the test's boundary.

$$\text{Sensitivity} = \frac{TP}{TP + FN} \times 100$$

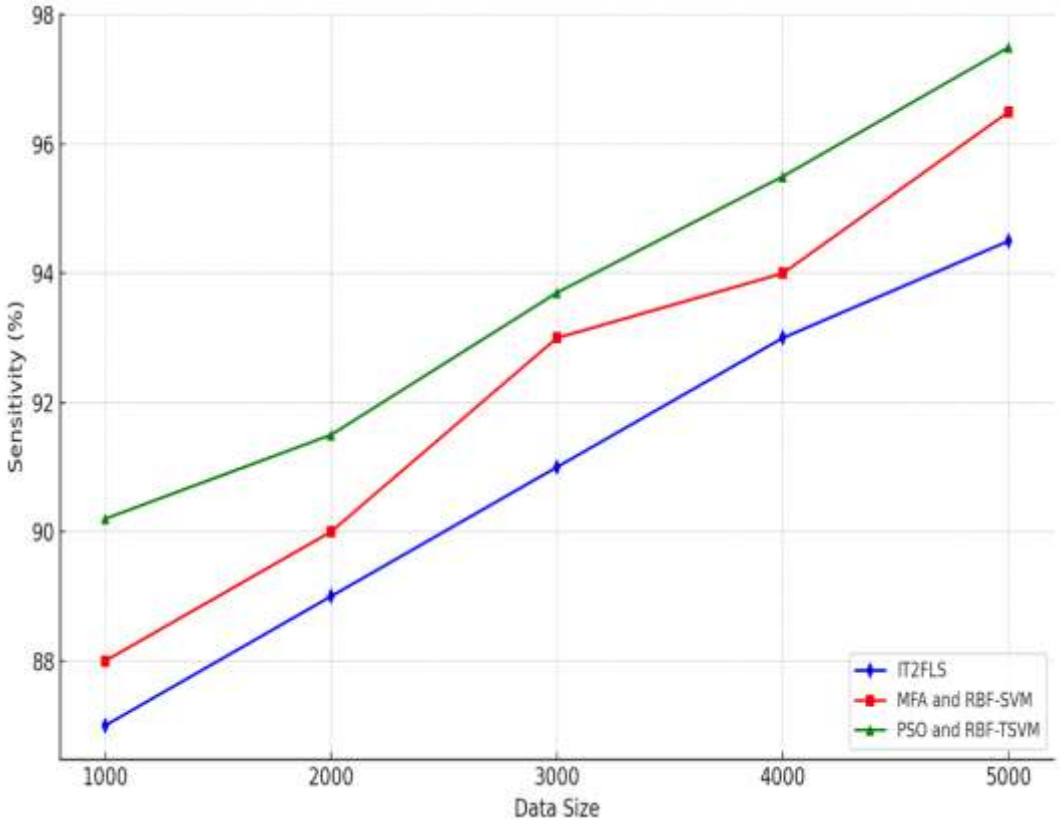

Figure 3.17. Results for Sensitivity

Figure 3.16 shows the sensitivity findings of the IT2FLS, RBF-SVM, and PSO-RBF-TSVM based classification methods, as well as the suggested PSO-RBF-TSVM based classification method. The sensitivity is shown on the Y-axis, while the dataset size is plotted on the X-axis. The suggested work employs the Z-Score method for normalization in an effort to improve the system's overall performance. Additionally, RBF-TSVM is used to accomplish effective categorization. The rate of true positives is enhanced. When compared to the current system, the suggested PSO and based RBF-

TSVM classification method exhibited good sensitivity findings across all dataset sizes.

(iii) Specificity-PSO–RBF-TSVM vs Baseline Models

The percentage of correctly identified negatives is the metric for specificity. The ability of the test to detect unfavorable outcomes is at the heart.

$$\text{Specificity} = \frac{TN}{TN + FP} \times 100$$

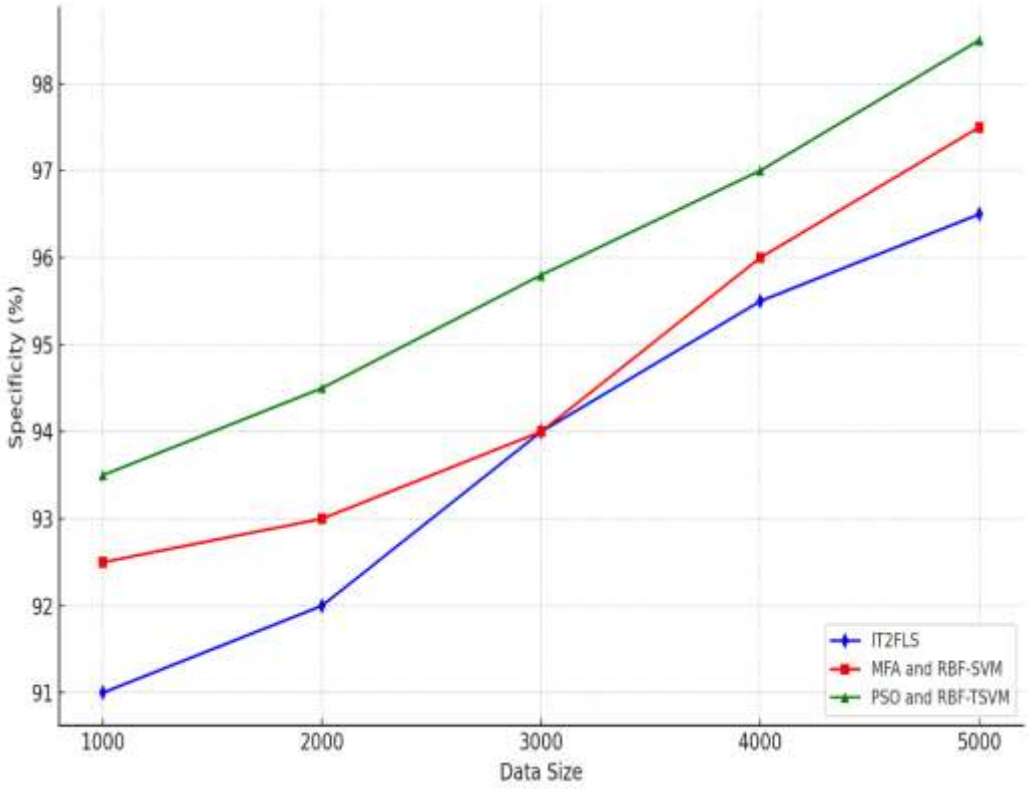

Figure 3.18. Results for Specificity

Figure 3.17 shows the specificity findings of the IT2FLS, RBF-SVM, and PSO-RBF-TSVM based classification methods, as well as the suggested PSO-RBF-TSVM based classification method. On the one hand, we have the dataset size (X-axis) and the specificity (Y-axis). The suggested PSO and based RBF-TSVM classification method outperformed the state-of-the-art system across all dataset sizes.

Table 3.7 Results of PSO-RBF-SVM and IT2FLS

Data Size	**Accuracy**		**Sensitivity**		**Specificity**	
	IT2FLS	PSO-RBF-TSVM	IT2FLS	PSO- RBF-TSVM	IT2FLS	PSO-RBF-TSVM
1000	91.5	93	87	90	91	93
2000	92	94	89	91.5	92	94
3000	94	95.5	91	93.5	93	95
4000	95	96.5	94	96	95	97

The table compares the performance of IT2FLS and PSO-RBF-TSVM across dataset sizes ranging from 1000 to 5000 instances using accuracy, sensitivity, and specificity. Although both models improve as the dataset grows, PSO-RBF-TSVM consistently achieves higher results. IT2FLS shows accuracy rising from 91.5% to 96%, sensitivity from 87% to 95%, and specificity from 91% to 96%. In contrast, PSO-RBF-TSVM increases accuracy from 93% to 98%, sensitivity from 90% to 97.5%, and specificity from 93% to 98%. These results indicate that PSO-RBF-TSVM delivers stronger and more reliable predictive performance than IT2FLS, especially with a larger dataset.

Table 3.8 Findings from the PSO-RBF-TSVM and MFA-RBF-SVM Models

Data Size	Accuracy		Sensitivity		Specificity	
	MFA and RBF-SVM	PSO-RBF-TSVM	MFA and RBF-SVM	PSO-RBF-TSVM	MFA and RBF-SVM	PSO-RBF-TSVM
1000	92	93	89	90	92	93
2000	93	94	91	91.5	93	94
3000	95	95.5	93	93.5	94	95
4000	96	96.5	95	96	96	97
5000	97	98	97	97.5	97	98

Table 3.9 Percentage wise Improvement of PSO-RBF-TSVM over IT2FLS

Data Size (Bytes)	Accuracy	Sensitivity	Specificity
1000	1.61	3.33	2.15
2000	2.12	2.73	2.12
3000	1.57	2.67	2.1
4000	1.55	2.08	2.06
5000	2.04	2.56	2.04

Table 3.10 Percentage wise Improvement of PSO-RBF-TSVM over MFA-RBF-SVM

Data Size	Accuracy	Sensitivity	Specificity
1000	1.07	1.11	1.07
2000	1.06	0.54	1.06
3000	0.52	0.53	1.05
4000	0.51	1.04	1.03
5000	1.02	0.51	1.02

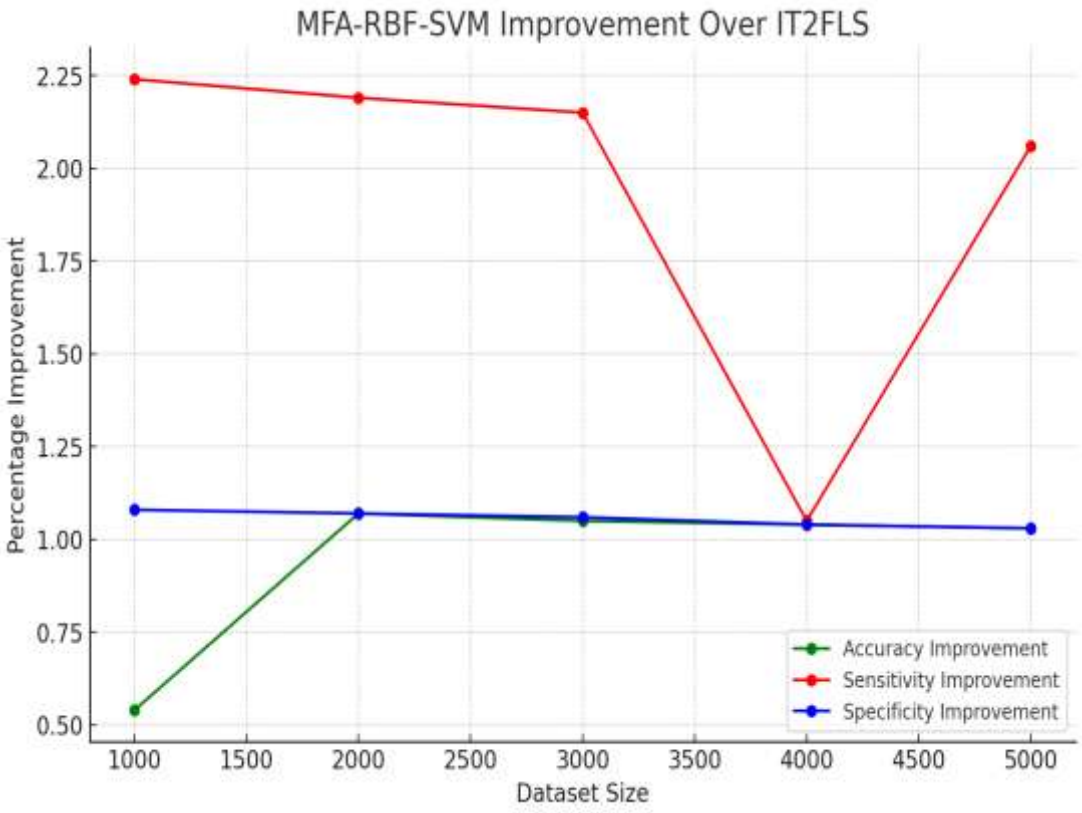

Figure 3.19 – Percentage Improvement of MFA–RBF-SVM Over IT2FLS Across Different Dataset Sizes

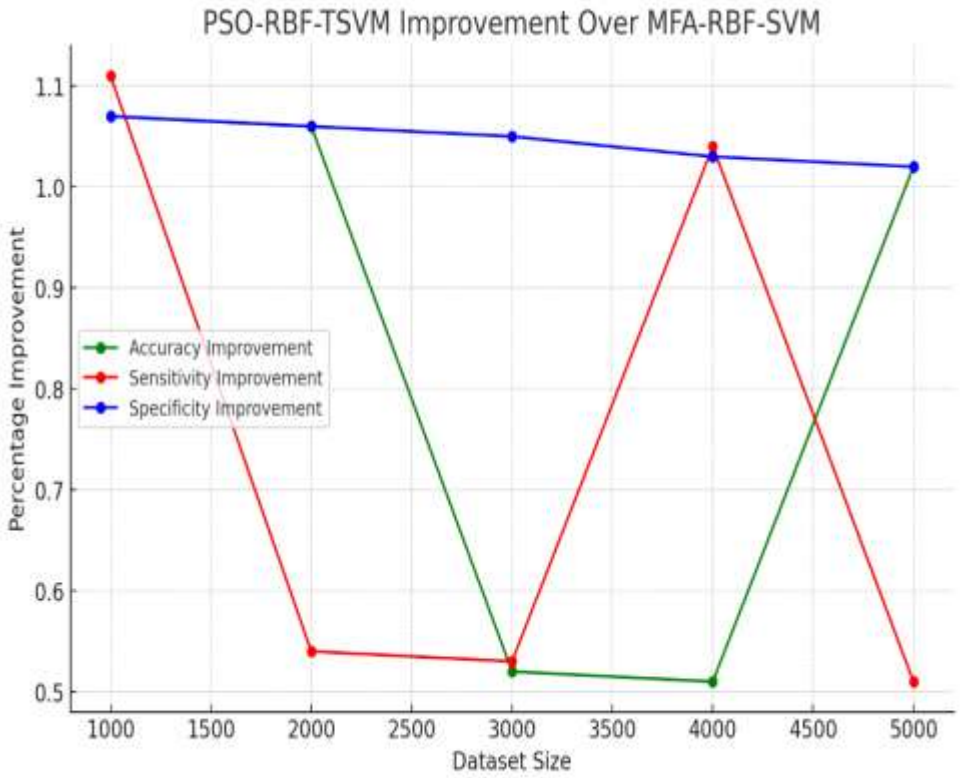

Figure 3.20 – Percentage Improvement of PSO–RBF-TSVM Over IT2FLS Across Different Dataset Sizes

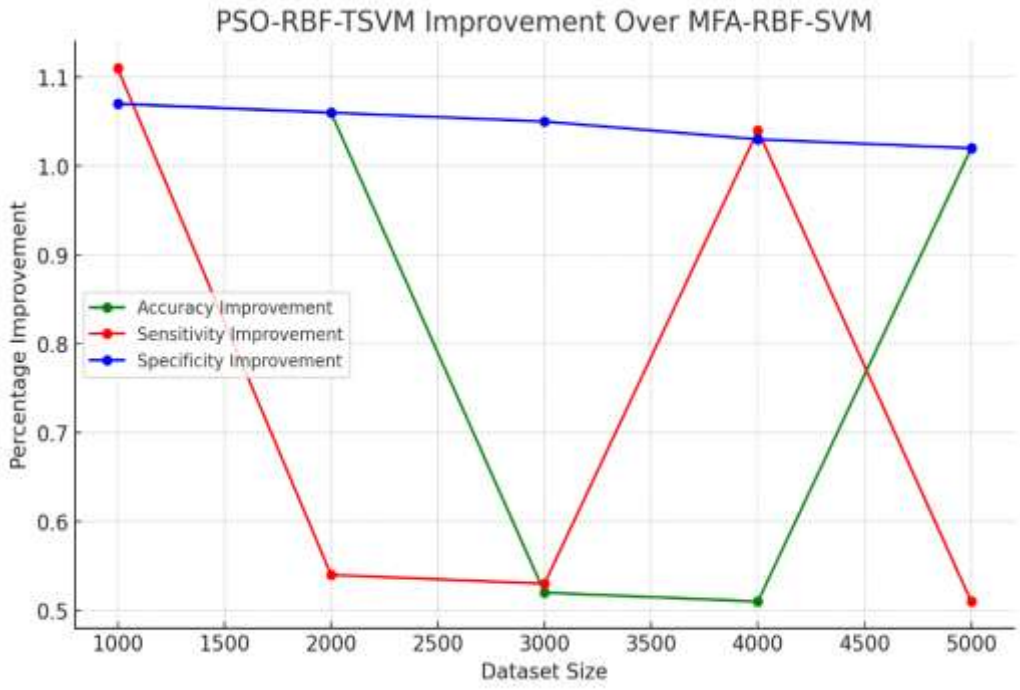

Figure 3.21 – Percentage Improvement of PSO–RBF-TSVM Over MFA–RBF-SVM

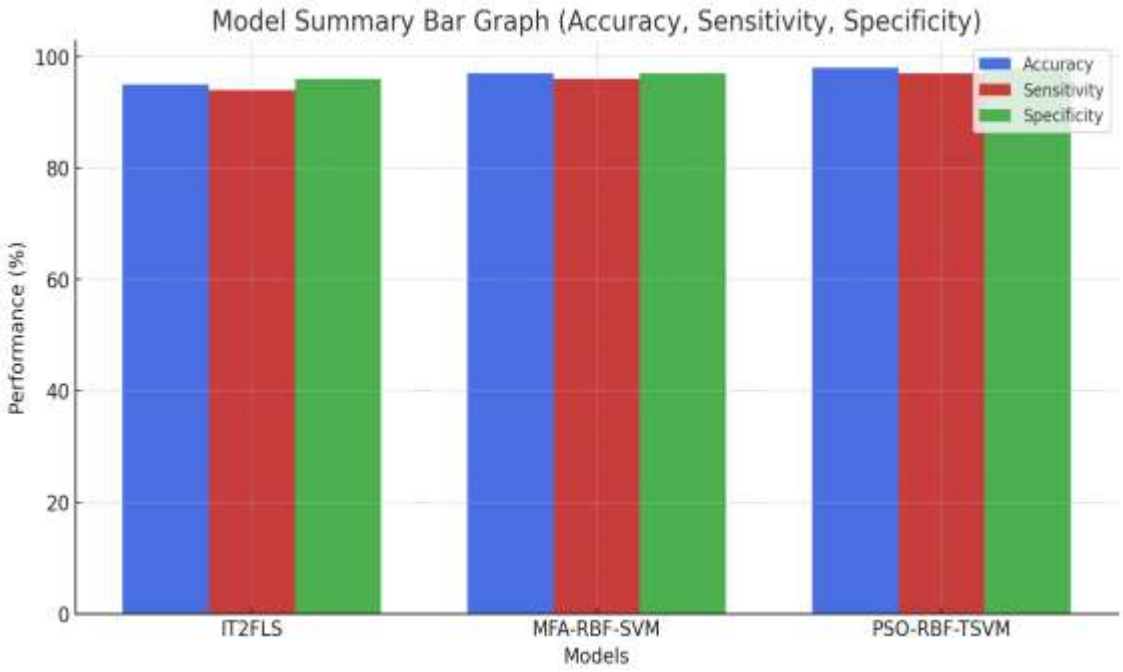

Figure 3.22– Summary Graph

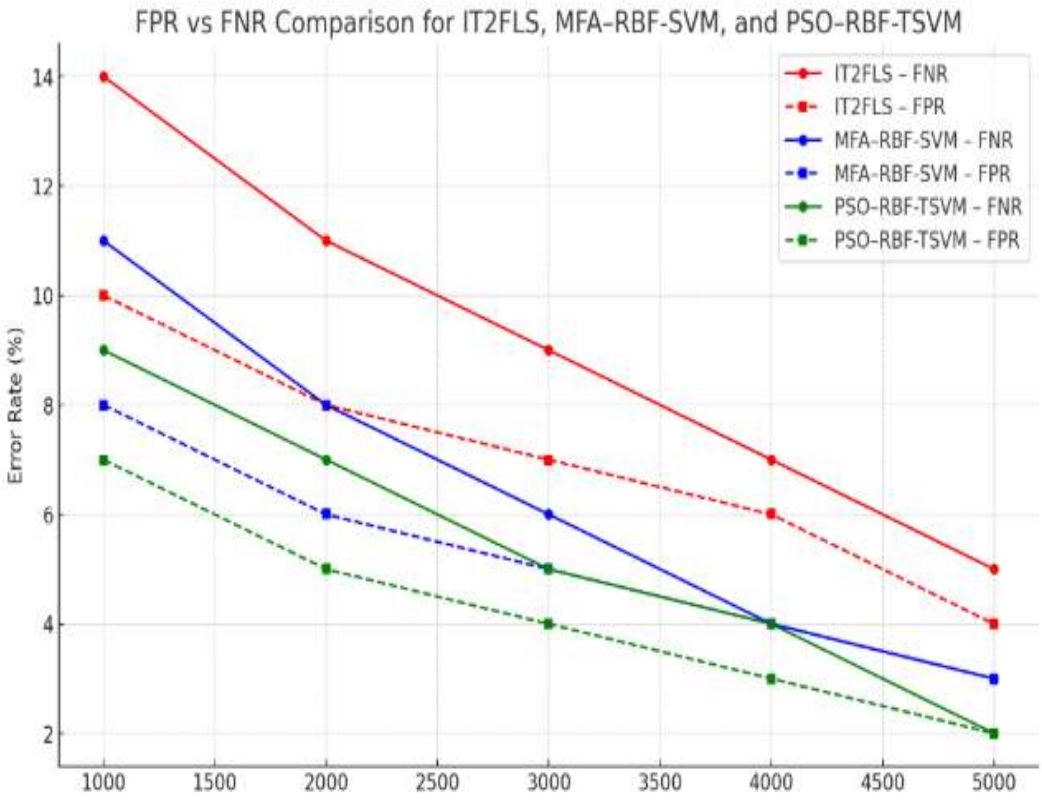

Figure 3.23 -FPR–FNR Comparison Across Models and Dataset Sizes

Table 3.11: overall average accuracy, sensitivity, specificity.

Model	Accuracy	Sensitivity	Specificity
IT2FLS	96%	95%	96%
MFA–RBF-SVM	97%	97%	97%
PSO–RBF-TSVM	98%	97.5%	98%

Table 312: FPR–FNR Comparison Across Models and Dataset Sizes

Dataset Size	IT2FLS FNR (%)	IT2FLS FPR (%)	MFA–RBF-SVM FNR (%)	MFA–RBF-SVM FPR (%)	PSO–RBF-TSVM FNR (%)	PSO–RBF-TSVM FPR (%)
1000	14	9	11	8	10	6
2000	11	8	8	6	7	4
3000	9	6	7	5	5	3
4000	7	5	5	4	4	3
5000	5	3	3	3	2	2

3.15 SUMMARY OF CHAPTER

Chapter 3 presented the complete methodological framework used to develop the proposed heart-disease prediction system, beginning with an overview of the research approach and a flowchart depicting the workflow. It described the dataset, its challenges, relevance, and the preprocessing steps applied, including handling missing values, treating outliers, scaling features, and encoding categorical data. The chapter explained the feature selection techniques (MFA, PSO-RS), classification methods (RBF-SVM, TSVM), and optimization strategies (MFA, PSO-RS, OCSO) employed to enhance model accuracy. It also justified the choice of the proposed hybrid models over the existing IT2FLS method and outlined the experimental setup, evaluation metrics, and statistical validation procedures. While Chapter 3 explains how the system is designed and implemented.

Chapter 4 is needed because it provides the empirical validation of the methodology developed in Chapter 3. While Chapter 3 explains the algorithms, preprocessing, and system design, Chapter 4 demonstrates *how well* these methods perform through detailed experimentation. It presents accuracy, sensitivity, and specificity results for OCSO, PSO, GA, and CSO; compares the proposed approach with other optimization methods; and includes graphical trends, percentage improvement calculations, and thorough discussion of classifier behavior. Chapter 4 also verifies the practical usefulness of the neutrosophic diagnosis model and confirms that the proposed OCSO-RBF-TSVM framework significantly improves prediction performance. Thus, Chapter 4 is essential for translating the methodology into measurable outcomes, proving its effectiveness, and validating the contributions of the study.

CHAPTER-4

A ROBUST SYSTEM USING ARTIFICIAL INTELLIGENCE AND SOFT COMPUTING TECHNIQUES FOR IDENTIFYING AND PREDICTING HEART DISEASES

In order to overcome the complexity and enhance performance compared to traditional methods of cardiac disease prediction, this chapter explains how Artificial Intelligence (AI) and soft computing approaches may be used. Collecting cardiac data in a real-world setting is the first step, followed by reducing data redundancy and improving data integrity. Zero-Score (Z-Score) is used to normalize the data. Following this, Multiple soft computing methods, including Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Crow Search Efficiency (CSO), and Opposition Based Crow Search Optimization (OCSO), are used to accomplish attribute reduction. Finally, we have RBF-TSVM, an acronym for Radial Basis Function-Transductive Support Vector Machines, which is a classifier. The findings demonstrate that the proposed OCSO technique outperforms the present method in terms of accuracy, sensitivity, and specificity.

The leading killers on a global scale are cardiovascular disorders. The mortality rate may be reduced if the condition can be detected early. Many new decision-making systems have recently been created, but their complexity prevents them from being used by healthcare practitioners. In order to accomplish these aims, a digestible neutrosophic clinical decision-making system is suggested, which would take 35 different characteristics into consideration. The most important parts of our suggested model are the neo-optimal approach, the inference engine, the rule building, the explainability, and the causality. To show how well our model works, we included an algorithm for calculating the risk of cardiovascular disease using a single-valued neutrosophic method. The model classifies heart disease severity on a scale from 1 to 5, using a Multi-Attribute Decision Making (MADM) approach that combines Interval-valued Trapezoidal Neutrosophic Numbers (IvTNN) and Weighted Aggregated Sum Product Assessment (WASPS).

Vulnerability is the most important and basic fact in the medical sector. Representing patients' emotions, physicians' opinions, and laboratory results correctly is next to impossible. No one in the field of clinical research has yet provided a satisfactory explanation for how diseases disrupt the body's usual processes. Many businesses, including the medical field in particular, provide decision-makers with a high degree of uncertainty. Important decisions need to be made by doctors swiftly and precisely. It is challenging for less experienced doctors and physicians to diagnose heart disease due to the wide variety of symptoms and pathologic characteristics.

4.1 EFFECTIVE METHODS FOR THE DIAGNOSIS OF HEART CONDITIONS WITH THE USE OF AI AND SOFT COMPUTING

Askarzadeh proposed a novel metaheuristics optimization method, which he called CSO, based on the crow's collective behavior in recent days. The concept of CSO is based on the notion of storing extra nutrients in hidden places and then reintroducing them at the critical moment. After other birds leave a spot where they've stashed food, the crow can smell it and takes it. Once the burglary is done, it admits to escape from becoming a prey even more.

If there are N crows in a flock, then it stands to reason that crow i will have location xi k at repetition k. The disguised spot where the food shadowed by crow i was kept. Crow explores the world in search of the best food source, m ik, in the exploration level. There are two potential outcomes for the CSO probing approach. The first is that the owner of the nourishment origin property, crow j, fails to distinguish between the burglar, crow i, and follows it. So, the crow of the thief lands on the crow of the owner, who has already conceded. The crow burglar's location alert method is defined by

$$X_i^{k+1} = X_i^k + r_i \times fl_i^k \times \left(m_j^k - X_i^k \right)$$

Here, ri is any positive integer in the interval [0, 1], and flik is the distance flown by crow i at iteration k.

As a second possibility, it's possible that owner crow j notices that burglar crow i is following it, and owner crow j decides to betray burglar crow i by going to a different investigating spot. A random location rearranges the position of crow i.

The following manifestation resolves the problem in CSA:

if $r_j \geq p_j^k$

update position by eqn.5.1

else update to random position

This is where pik represents the probability of crow j's awareness at repetition k and rj is any integer in the interval [0, 1].

The parameter fl_i^k is crucial in determining the best answer on a worldwide scale, even with just a tiny amount of fl_i^k suggestions for the local lowest even when huge sums result in a worldwide minimum, The results of are shown in Figures 4.1 and 4.2 fl_i^k as part of the search process.

At repetition k, the group's location is rearranged according to Eq.(5). At the same time, The appropriateness task is evaluated first-hand. At iteration k, the completed suitability task is connected to the one indicated before, and the group's location is placed to alert.

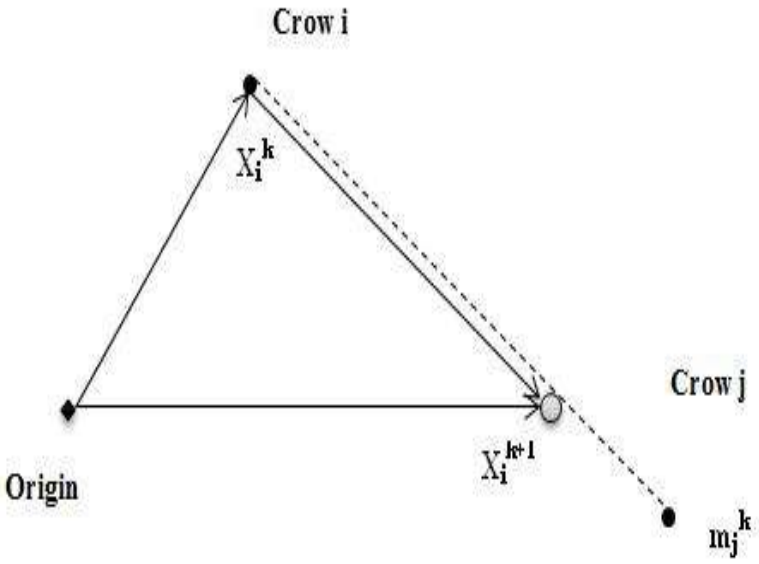

Figure 4.1. Parameter fl<1

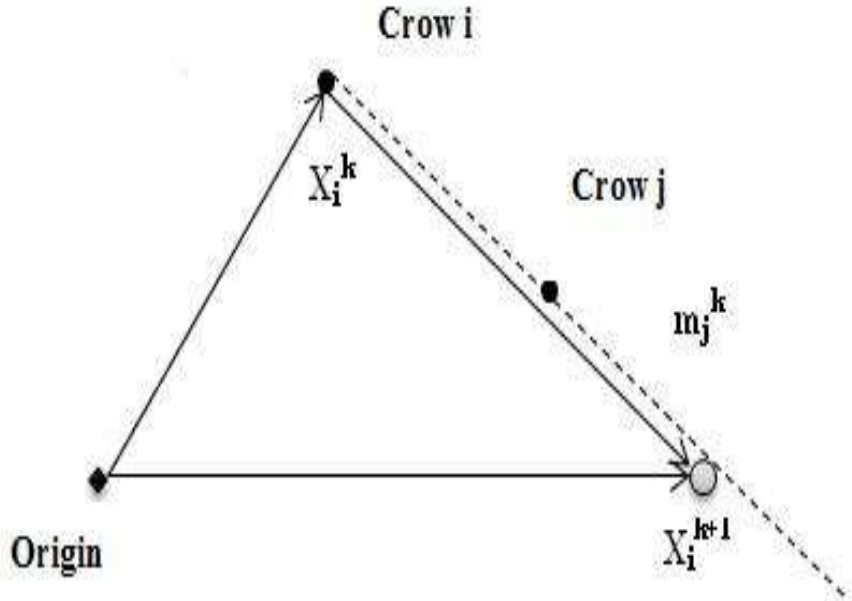

Figure 4.2 Parameter fl>1

CROW SEARCH ALGORITHM

Randomly initialize the position of a flock of N crows in the search space
Evaluate the position of the crows
Initialize the memory of each crow

while iter < itermax do
  for i = 1 to N (all N crows of the flock) do
    Randomly choose one of the crows to follow (for example, j)
    Define an awareness probability (AP)

    if rand >= AP then
      x_i^(iter+1) = x_i^iter + fl * (m_j^iter - x_i^iter)
    else
      x_i^(iter+1) = a random position in the search space
    end if
  end for

  Check the feasibility of new positions
  Evaluate the new position of the crows
  Update the memory of the crows
end while

- **OCSO implementation for optimization**

Here, we'll assume that you know how to implement OCSO step-by-step.

Step 1: Setup and changeable constraints

We classify the inflation problem, resolution changes, and constraints. The OCSO's tunable parameters (flock size, N), which dictate the number of repeats ($iter_{max}$), At that point, the flying distance (fl) and the awareness probability (AP) are respected.

To demonstrate the feasibility of opposition-based learning, Hamid R. Tizhoosh introduces it and uses it to solve a few optimization issues. The main idea behind this

topology is to find an optimum solution that is just around the corner by creating an opposition-based solution for the first developed random solution. To compare the final result to the original, opposing answer, the following function is used.

$$Op_i = x + y - I_i$$

I_i $I_1, I_2, I_3 \ldots NP$, the above equation (4) proposes generating a counter-intuitive answer, with x and y being the extremes of the possible range, and these starting solutions are created at random.

Step 2: Get the crows' nest and memories ready.

In a d-magnitude exploring region, N crows are placed at random as group members. With d being the integer of resolution change, each crow confirms a likely solution to the problem.

$$Crows = \begin{bmatrix} X_1^1 & X_2^1 \_\_ X_d^1 \\ X_1^2 & X_2^2 \_\_ X_d^2 \\ - & - & - \\ - & - & - \\ - & - & - \\ X_1^N & X_2^N \_\_ X_d^N \end{bmatrix} \quad (5.4$$

Set each crow's memory to its initial state. Consequently, the crows play no role in the first recapitulation. Their nutritional secretion at their early locales is pretty typical.

$$Memory = \begin{bmatrix} m_1^1 & m_2^1 \ldots\ldots m_d^1 \\ m_1^2 & m_2^2 \ldots\ldots m_d^2 \\ . & . & . \\ . & . & . \\ . & . & . \\ m_1^N & m_2^N \ldots\ldots m_d^N \end{bmatrix}$$

Step 3: Estimate Fitness function

By inserting the verdict flexible miles into the impartial job, we may assess the excellence of every crow's place.

Using the correctness of characteristic reduction without insignificant reductions, these wellness capabilities may be evaluated. Applying a health task that considers both trait reduction (nature of guess characterization) and insignificant property reduction is crucial for finding the optimum negligible characteristic decrease. Thus, this inquiry is linked to a wellness task, as indicated in Eq.

$$Fitness\ (X) = \frac{m - |X|}{m} + \frac{n|R|\gamma_X(D)}{m\Gamma}$$

Here $m=|C|$, $|U|$,$\gamma\_x\ (D)$ that arrangement has is called its nature. R is recorded using a regulated fast reduction computation, and it is known as a reduce of scenario trait C.

For the characteristic reduction job, this formula implies that the length of the qualities subset, |X|, and the course of action value, γ_x (D), are of distinct relevance.

Step 4: Design a fresh setting

Crows create new territory at the investigation site while investigators conduct follow-up inquiries: Think about a crow—I need one to create a new space. This crow (mej) selects one of the group crows at random, so you may follow its trail to find out what happened to the food it was hiding. The vast majority of crows follow this pattern.

Step 5: Investigate potential new employment opportunities

Crow does something unexpected; it remains where it is instead of flying to a newly-created spot.

Step 6: Assess the suitability of a new site

For each crow's unique location, we evaluate their suitability task rate.

Step 7: Bring memories that are current

In their subsequent searches, the crows update their memory:

$$me^{j,iter+1} = \begin{cases} X^{i,iter+1} & f(X^{i,iter+1}) \ \ is \ better than \ f(me^{j,iter}) \\ me^{j,iter} & O.W \end{cases}$$

The objective task rate is deduced by f(.).

Clearly, if a crow's suitability rate in its new site is higher than its suitability rate at its old one, then.

Step 8: Check termination criterion

The fourth through seventh stages are tedious until itermax is reached. Recalling that the closure principle is the solution to the inflation issue is a good place to start when thinking about the fair task rate.

4.2 CARDIOLOGICAL NEPHROSOPHIC CLINICAL DECISION-MAKING SYSTEM

The doctor makes a call based on past judgments made for patients in comparable scenarios and an estimation of the patient's actual examination findings. The knowledge and skill of a doctor make this feasible, but because there are so many clinical, behavioural and physiological factors to consider, this task becomes extremely time-consuming. Therefore, there is an urgent need for a precise and intelligent system that can detect patients whose conditions, symptoms or risk patterns are similar or identical. Although ML algorithms will play an increasingly important role in illness prediction, traditional ML approaches were never designed to handle the level of uncertainty, inconsistency and incompleteness present in real clinical data. Many of the learning problems are expressed using contradictory, imprecise or missing facts, and as a result, the performance of conventional ML models becomes insufficient when applied to medical datasets. In addition, data preparation itself is computationally intensive and prone to errors arising from data collection, feature extraction and reporting. When

inaccurate or incomplete data is provided, ML systems face genuine learning obstacles. These limitations create a clear research gap: existing prediction models cannot reliably handle uncertainty, and therefore their performance degrades when applied to real-world cardiac datasets such as ours.

To address this, Single-Valued Neutrosophic Sets (SVNs) provide a mathematical framework for modelling inaccurate, vague and indeterminate information. Unlike traditional ML methods that rely solely on crisp input values, neutrosophic learning algorithms are capable of manipulating data that contains truth, falsity and indeterminacy simultaneously, making them highly suitable for complex medical environments. As an extension of the fuzzy system proposed by Smarandache (1995), neutrosophic statistics (NS) directly address uncertainty that fuzzy logic alone cannot capture. In neutrosophic classifiers, every piece of data includes three components—true, false and ambiguous—allowing the model to preserve uncertainty rather than forcing a premature decision. This helps clinicians obtain more realistic predictions and significantly reduces uncertainty in medical decision-making. NS has been widely adopted in various fields, including medicine, physics, computer science and engineering, demonstrating its robustness for complex domains. In the proposed chapter, this framework is applied to classify heart disease data more efficiently, and the cardiac dataset is evaluated using a neutrosophic diagnostic method. The model diagram for neutrosophic-based disease prediction (Figure 4.3) shows how neutrosophic reasoning determines the relative importance of each feature to forecast cardiovascular disease.

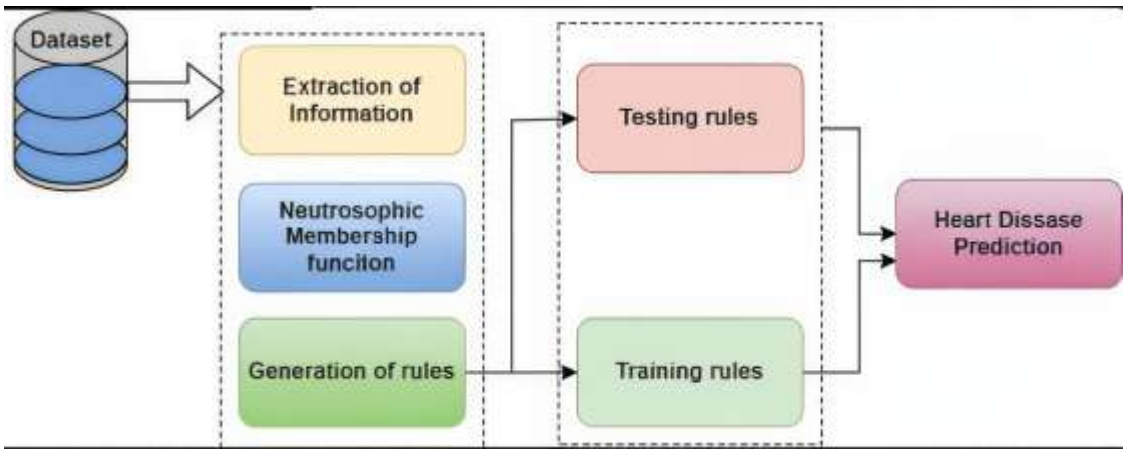

Figure 4.3 Model Diagram for Neutrosophic-Based Disease Prediction

To implement this approach, relevant data is first extracted, followed by defining neutrosophic membership functions and establishing the criteria for illness prediction. The system is then trained and tested for accurate disease categorisation. Although several ML and DL methods exist for CVD prediction, most function effectively only in controlled clinical settings. However, modern healthcare requires prediction systems that work in dynamic real-time monitoring environments. This chapter therefore integrates neural network techniques with neutrosophic sets to build a more reliable prediction model suitable for real-time cardiac risk assessment. The goal is to detect the onset of a heart attack and alert physicians in advance. This requires identifying complex interactions among dependent and independent variables, for which classical methods are insufficient.

This decision-making process is designed to support medical professionals, but clinical staff often hesitate to trust computational systems due to lack of transparency. Neutrosophic reasoning addresses this by providing interpretable outputs that enhance trust and promote adoption. The proposed research is notable because it introduces a reliable technique for multi-classifying cardiac data and comparing confusion matrix results with other fuzzy and soft computing algorithms. Although diagnostic tests in classical statistics assume that each datapoint is precisely known, real clinical datasets frequently contain vague or ambiguous observations. In such cases, neutrosophic statistics provide a more realistic analytical framework. Health datasets contain significant ambiguity that, if improperly managed, can lead to misdiagnosis. Furthermore, increased system complexity often makes clinicians reluctant to adopt decision-support tools. By using neutrosophic reasoning, the proposed model overcomes these limitations and provides a robust prediction framework for CVD. Thus, the study effectively addresses the need for a more accurate, uncertainty-aware, real-time prediction system for heart disease, bridging a critical gap not addressed by existing ML or DL methods.

4.3 CATEGORIES OF CVD

Various types of cardiovascular disorders may be classified. Each group is described in Table 4.1

Table 4.1 Several Types of Cardiovascular Diseases with Their Description

Sl. No	Category of CVD	Disease Value	Description
1	Coronary artery disease	D1	A CAD is a constriction of the coronary arteries. Fat and cholesterol in the blood vessel called atherosclerosis. These plaques may block the artery, preventing blood from reaching the heart muscle.
2	Heart arrhythmia	D5	It is called arrhythmia which means the heartbeat rate is too fast or slow or the interval of the heartbeat becomes irregular. The heart is one kind of electrical system of the human body, which handles the heartbeat and circulates blood throughout the body. If anything goes wrong in the system then the heart rhythm becomes abnormal.
3	Peripheral artery disease	D6	Narrow arteries are a frequent cardiovascular problem that restricts blood flow to the organs. The spheres do not receive enough blood flow as the body starts to develop PAD.
4	Heart valve disease	D7	It is one kind of hereditary. In adults, it can also be caused by a variety of factors and conditions, including infections and other cardiac problems. Regression of heart valves can occur.
5	Heart failure	D8	Heart failure occurs in the human body when the muscles of the heart do not pump the blood as required. Due to such a problem, the heart gradually weakens or tightens and could not fill and pump effectively.

By feeding the suggested system 35 different symptoms, we may calculate the risk of different CVDs. The specifics are included in the section that follows table 4. The binary association between symptoms and CVDs is shown in 2 below, with 'ι' representing binary 1 and 'χ' representing binary 0. To find out how accurate the final cardiovascular disease outline is, this table will be helpful.

Table 4.2 Binary Correlation between Symptoms and CVDs Type Of CVD

Symptom ID	**D1**	**D2**	**D3**	**D4**	**D5**	**D6**	**D7**	**D8**
Sym-I	J	X	X	J	X	J	X	J
Sym-II	J	J	X	J	X	X	X	J
Sym-III	X	X	X	X	X	J	J	J
Sym-IV	J	X	J	J	X	J	X	J
Sym-V	J	X	X	X	J	X	J	J
Sym-VI	J	X	J	X	J	J	J	X
Sym-VII	J	X	X	J	J	J	X	J
Sym-VIII	J	J	X	J	J	J	X	J
Sym-IX	X	J	J	X	X	J	X	J
Sym-X	J	J	X	J	J	J	X	X
Sym-XI	J	X	J	J	X	X	J	X
Sym-XII	X	J	X	J	X	J	X	X
Sym-XIII	J	J	J	J	X	X	J	J
Sym-XIV	X	X	J	X	X	J	X	J
Sym-XY	X	X	J	J	X	J	X	J
Sym-XYI	J	X	X	J	J	X	J	J
Sym-XYII	X	J	X	X	J	J	X	X
Sym-XYIII	X	J	J	J	X	J	X	J
Sym-XIX	X	X	J	J	J	X	X	J
Sym-XX	X	X	X	J	J	X	J	X
Sym-XXI	J	J	X	J	X	J	J	J
Sym-XXII	X	X	J	X	X	J	J	J

Sym-XXIII	X	X	J	X	J	X	J	X
Sym-XXIV	J	J	J	X	X	X	J	J
Sym-XXV	X	J	X	J	X	X	X	X
Sym-XXVI	X	J	X	J	X	X	J	J
Sym-XXVII	X	J	X	X	J	J	J	J
Sym-XXVIII	X	J	X	X	J	J	X	J
Sym-XXIX	X	X	X	X	X	X	J	X
Sym-XXX	X	X	X	X	X	X	X	X
Sym-XXXI	X	X	X	X	X	X	X	J
Sym-XXXII	J	J	J	J	J	X	J	X
Sym-XXXIII	J	J	X	X	X	X	J	X
Sym-XXXIV	J	J	J	J	X	J	X	X
Sym-XXXV	X	X	J	X	J	J	X	J

4.4 PRELIMINARIES OF NEUTROSOPHIC SETS

The NS concept is derived from neutrosophy, a new path of philosophy [129]. “A set of points ξ with the general element in ξ represented by . Then a neutrosophic set α in ξ is defined set of membership functions T , I andF . A subset of these membership functions includes the truth membership function (TMF) and the indeterminacy membership function (IMF).] 0̄, 1 [, that represents T : ξ →] 0̄, 1 [; I : ξ →] 0̄, 1 [F ; ξ →] 0̄, 1 [.”

An SVN includes a non-empty set Xon. The TFM defines Y as T : Y → [0,1], the IMF defined as I : Y → [0,1] and FMF defined as F : Y → [0,1]. Where

S={<a,T (a), I (a), F (a) > |a □ Y}[7] .

An SVN number, denoted as X, is

X([M1 , M2 , M3 , M4); ρ]; ([M1 , M2 , M3 , M4); σ],([M1 , M2 , M3 , M4); ω]wh ere ρ, σ and ω □ [0,1]. The TMF, IMF μ : R [0, ρ], (ν): R → [σ, 1] and(λ): R →[ω, 1] respectively are defined in equations

$$\mu_M(Q) = \begin{cases} \mu_{M1}(Q), & \text{ifm}_1 \le Q \le n_1, \\ \rho, & \text{if } n_1 \le Q \le o_1, \\ \mu A_\mu(Q), & \text{ifo}_1 \le Q \le q_1, \\ 0, & \text{otherwise.} \end{cases}$$

$$\nu_M(Q) = \begin{cases} \nu_{M1}(Q), & \text{ifm}_2 \le Q \le n_2, \\ \rho, & \text{ifn}_2 \le Q \le o_2, \\ \nu A_\mu(Q), & \text{ifo}_2 \le Q \le q_2, \\ 1, & \text{otherwise.} \end{cases}$$

$$\lambda_M(Q) = \begin{cases} \lambda_{M1}(Q), & \text{ifm}_3 \le Q \le n_3, \\ \rho, & \text{ifn}_3 \le Q \le o_3, \\ \lambda A_\mu(Q), & \text{ifo}_3 \le Q \le q_3, \\ 1, & \text{otherwise.} \end{cases}$$

The ∩of two SVNS can be expressed as $X_3 = X_1 \cap X_2$:

$$A_{s3}(Q) = \min\left(A_{s1}(Q), A_{s2}(Q)\right),$$
$$B_{s3}(Q) = \max\left(B_{s1}(Q), B_{s2}(Q)\right)$$
$$C_{s3}(Q) = \max\left(C_{s1}(Q), C_{s2}(Q)\right) \text{ for all } Q \text{ in } R$$

The □ of two SVNS can be expressed as $X_3 = X_1 \cap X_2$:

For SVNS truth membership functions, the formula for mathematical computation is given by the equation above. Two sets of SVNS may be defined by the union and intersection in Equations.

$$A_{s3}(Q) = \max\left(A_{s1}(Q), A_{s2}(Q)\right), B_{s3}(Q) = \min\left(B_{s1}(Q), B_{s2}(Q)\right), C_{s3}(Q) = \min\left(C_{s1}(Q), C_{s2}(Q)\right) \text{ for all } Q \text{ in } R$$

- For SVNS truth membership functions, the formula for mathematical computation is given by the equation above. Two sets of SVNS may be defined by the union and intersection in Equations.

- **Illustrating Neutrosophic Technique for Heart Disease Decision-Making System**

Decisions are being made via SVNs. We have used explainable AI methods to make it simpler for doctors to grasp. We make sure that every module of the system has an explanation section. You may learn more about that module's inner workings in the interpretation case. Its time complexity may be calculated using the suggested approach. The algorithm assesses the risk of various cardiovascular illnesses given in Table 4.3 using 35 different types of factors as input. The schematic representation of the proposed system is presented in Figure 4.4. NL tool that is integrated into the proposed system.

Table 4.3 List of Input Variables

Sl. No	Symptom ID & Variable	Range Value
1	Sym-I: 'Gen'	01-Feb
2	Sym-II: 'Age'	0–110
3	Sym-III: 'Genetic Nature'	01-Feb
4	Sym-IV: 'Smoking'	01-Feb
5	Sym-V: 'Systolic BP'	90–150 (mm Hg)
6	Sym-VI: 'Cholesterol'	100–400 (mg/dL)
7	Sym-VII: 'Diabetes'	68–300 (mg/dL)
8	Sym-VIII: 'BMI'	10–40 (kg/m^2)
9	Sym-IX: 'Depression'	0–2
10	Sym-X: 'Unhealthy Diet'	0–2

11	Sym-XI: ‘Metabolic Disorder’	01-Feb
12	Sym-XII: ‘Physical Inactivity’	01-Feb
13	Sym-XIII: ‘Pre-eclampsia’	01-Feb
14	Sym-XIV: ‘Rheumatoid arthritis’	01-Feb
15	Sym-XV: ‘Consumption of Coffee’	01-Feb
16	Sym-XVI: ‘Pregnancy’	01-Feb
17	Sym-XVII: ‘Rubella’	01-Feb
18	Sym-XVIII: ‘Usage of Drugs’	01-Feb
19	Sym-XIX: ‘Tobacco’	01-Feb
20	Sym-XX: ‘Alcohol’	01-Feb
21	Sym-XXI: ‘Heart problem’	01-Feb
22	Sym-XXII: ‘Past injury’	01-Feb
23	Sym-XXIII: ‘Thyroid’	01-Feb
24	Sym-XXIV: ‘Sleep apnea’	01-Feb
25	Sym-XXV: ‘Atrial branching’	01-Feb
26	Sym-XXVI: ‘Past heart functioning history’	01-Feb
27	Sym-XXVII: ‘Infection’	01-Feb
28	Sym-XXVIII: ‘Level of Homocysteine’	0–2
29	Sym-XXIX: ‘Pericardial Cysts’	01-Feb
30	Sym-XXX: ‘Marfan’	01-Feb

31	Sym-XXXI: ‘Syphilis’	01-Feb
32	Sym-XXXII: ‘Inflammation’	01-Feb
33	Sym-XXXIII: ‘Clots’	01-Feb
34	Sym-XXXIV: ‘Cancer’	01-Feb
35	Sym-XXXV: ‘Electrolyte disparity’	01-Feb

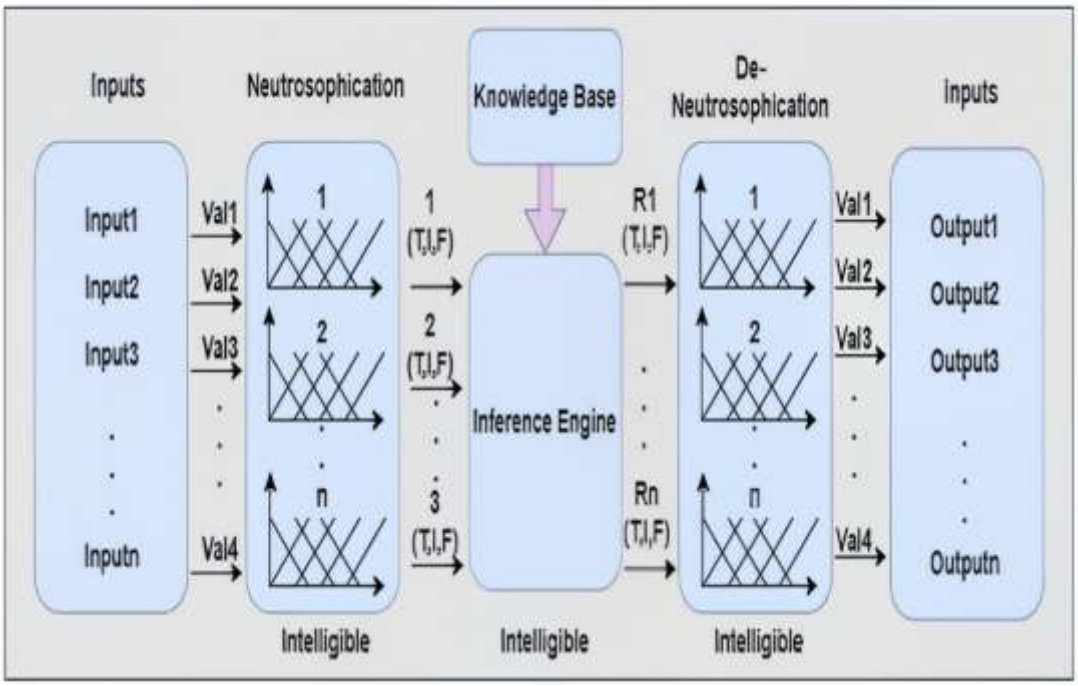

Figure 4.4 Block Diagram for the Proposed System

This diagram illustrates a Neutrosophic Decision-Making System, which is commonly used in intelligent healthcare applications to handle uncertainty and imprecision in medical data. The process begins with raw inputs, such as patient health information like age, cholesterol level, or blood pressure. These inputs are then passed through a stage called neutrosophication, where each value is transformed into a neutrosophic set, representing three components: Truth (T), Indeterminacy (I), and Falsehood (F). This approach allows the system to understand not just whether something is true or false, but also how uncertain the information is.

Next, the neutrosophic data enters the inference engine, which works with a knowledge base. This part of the system uses rules and medical expertise to process the inputs and make decisions, taking into account all three neutrosophic values (T, I, F). After decision-making, the system performs de-neutrosophication to convert the complex neutrosophic outputs back into clear, actionable results.

Finally, the system provides outputs—these could be predictions, diagnoses, or treatment suggestions. The entire model is particularly useful in healthcare for making intelligent decisions when patient data is vague, incomplete, or conflicting.

NL is a logic where each hypothesis is evaluated according to its probability in a certain subset, I, its degree of uncertainty, and F. Figure 4.5 shows the suggested model's flow diagram.

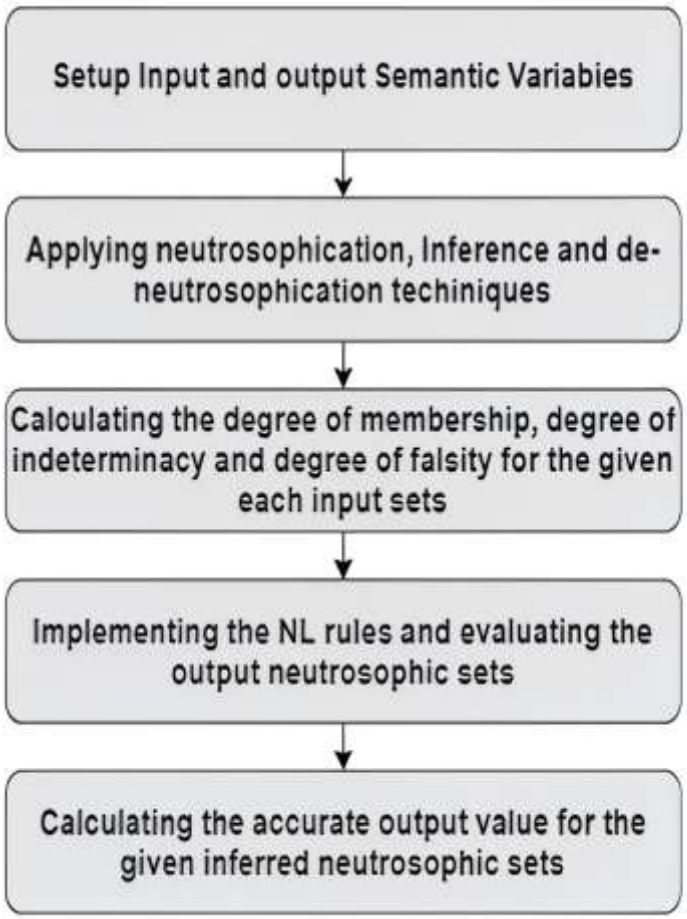

Figure 4.5 Visualization of the Suggested Mode

The first algorithm for illness risk assessment is this:

Algorithm 1:Calculating CVD risk using the SVN technique

Step 1: Acknowledge each of the 35 factors

Step 2: The degree of membership is determined using TMF, IMF, and FMF.

Step 3: Use rule construction to determine the strength of specific rules.

$$T_{(x)} = \min(\mu_{gender}(x), \mu_{age}(x), \dots \dots \mu_{EI}(x)),$$
$$I_{(x)} = \min(\mu_{gender}(x), \mu_{age}(x), \dots \dots \mu_{EI}(x)),$$
$$F_{(x)} = \max(\mu_{gender}(x), \mu_{age}(x), \dots \dots \mu_{EI}(x)),$$

Step 4: Use a de-neutrosophication method to determine the end result.

$$tot = \frac{(A + 2B + C + D + 2E + F + G + 2H + I)}{12}$$

Here we are using the new de-neutrosophication formula to the truth membership points (A, B, C), indeterminacy membership points (D, E, F), and falsity membership points (G, H, I):

$$tot = \frac{\sum_{i=1}^{n} x^i((3 - (\mu_x i + \nu_x i + \lambda_x i)) + \mu_x i \nu_x i \lambda_x i)}{\sum_{i=1}^{n}((3 - (\mu_x i + \nu_x i + \lambda_x i)) + (\mu_x i \nu_x i \lambda_x i))}$$

Step 5: Find the potential danger of each illness.

Step 6: Maximum value extraction from the output

Step 7: Write up As a consequence, the maximum risk

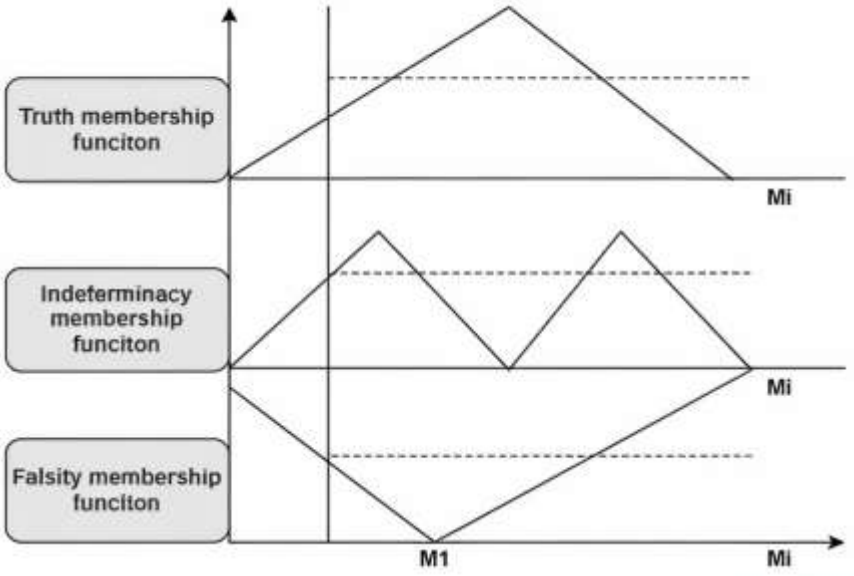

Figure 4.6 Neutrosophic Technique

The neutrosophic method converts data into language ideas by instituting statistical analysis. The neutrosophic approach takes three inputs—TMF, IMF, and FMF—as shown in Figure 4.6. Here are the results of the mathematical calculations:

$$\mu_{low}(x) = \begin{cases} \frac{110-p}{20}, \text{if } x \in [90,110], \\ 0, \quad \text{otherwise.} \end{cases}$$

$$\mu_{medium}(x) = \begin{cases} \frac{x-9}{25} \text{ if } x \in [90,120] \\ \frac{138-x}{18} \; x \in [90,138] \\ 0, \quad \text{otherwise.} \end{cases}$$

$$\mu_{high}(x) = \begin{cases} \frac{x-13}{10}, \text{if } x \in [135,145], \\ 0, \quad \text{otherwise.} \end{cases}$$

$$\nu_{low}(x) = \begin{cases} \frac{x-95}{20}, \text{if } x \in [95,135], \\ 1, \quad \text{otherwise.} \end{cases}$$

$$\nu_{medium}(x) = \begin{cases} \frac{125-x}{27} \text{ if } x \in [98,125] \\ \frac{x-127}{17} \; x \in [125,142] \\ 1, \quad \text{otherwise.} \end{cases}$$

$$\nu_{high}(x) = \begin{cases} \frac{145-x}{5} & \text{if } x \in [140,145] \\ 1 & x \in [145,150] \\ 1, & \text{otherwise.} \end{cases}$$

$$\lambda_{low}(x) = \begin{cases} \frac{x-92}{16}, & \text{if } x \in [92,108], \\ 1, & \text{otherwise.} \end{cases}$$

$$\lambda_{medium}(x) = \begin{cases} \frac{120-x}{23} & \text{if } x \in [97,120 \\ \frac{x-120}{15} & x \in [120-13 \\ 1, & \text{otherwise.} \end{cases}$$

$$\lambda_{high}(x) = \begin{cases} \frac{145-x}{7}, & \text{if } x \in [138,145], \\ 1, & \text{otherwise.} \end{cases}$$

$$\lambda \frac{yes}{female}(x) = 0$$

$$\lambda \frac{no}{male}(x) = 0$$

$$\mu \frac{yes}{female}(x) = 1$$

$$\mu \frac{no}{male}(x) = 1$$

$$\nu \frac{yes}{female}(x) = 0$$

$$\nu \frac{no}{male}(x) = 0$$

To determine low, medium, and high blood pressure, the mathematical calculation is shown.

Construction For Identifying the Disease Level

Ruling construction is the central component of decision-making systems. By using a rule foundation on knowledge-based data, this component generates new data. One

part of it is the rules that convey the user's choices. The rules are specified using If-Then control statements. The following are descriptions of several rules:

Rule 1:

IF: ((This person is male, they smoke, they use tobacco, they are physically inactive, they have inflammation, they clot, they have a predisposition to these conditions, and they are in their middle years of life.)) And (cholesterol=high, BP=high, diabetes=high, unhealthy diet=frequent, BMI=medium, homocysteine level=low, depression=high) And (rheumatoid alcohol=no, arthritis=no, rubella=no, metabolic disorder=no,drugs=no, thyroid=no, pre-eclampsia=no,coffee consumption=no, previous surgery =no, heart defect=no, pericardial cysts=no, marfan syndrome=no, syphilis=no , atrial fibrillation=no, thyroid=no, sleep apnea=no, pregnancy=no, infection=no, cancer=no, heart history=no, electrolyte imbalance=no))

Then,

extremely high rates of heart attacks, congenital heart defects, peripheral artery disease, arterial disease, arrhythmias, valve diseases, cardiomyopathy, and heart failure; low rates of heart arrhythmias; extremely low rates of heart valve disease; and low rates of heart failure.

Rule 2:

IF: ((gen=male or gen=female) And (Age=middle-age or age=old) And (smoking=no, genetic disposition=yes, tobacco=no , physical inactivity=no,) And (inflammation=yes, clots=yes , heart history=yes, alcohol=yes , cancer=yes, infection=yes, atrial fibrillation=yes, cholesterol=low, BP=medium, diabetes=medium, unhealthy diet=low, BMI=medium, homocysteine level=low, depression=low) And (rheumatoid arthritis=no, heart defect=no, rubella=no, previous surgery =no, drugs=yes, metabolic disorder=no, pre-eclampsia=no, coffee consumption=no, pericardial cysts=no, electrolyte imbalance=no, thyroid=no, syphilis=no ,thyroid=no, sleep apnea=no, marfan syndrome=no, pregnancy=no))

Then,

Cardiomyopathy ranks medium, heart failure ranks medium-low, cardiac arrhythmia ranks low, heart attacks rank low, The following cardiac conditions rate low: peripheral artery disease, congenital heart disease, and heart valve disease.

[End of IF]

4.5 ASYSTEM FOR NEUTROSOPHIC CLINICAL DECISIONS

The risk of different CVDs may be evaluated using the suggested approach and 35 different symptoms. The suggested neutrosophic clinical decision-is shown in Figure 4.7. What follows is a discussion of the specifics.

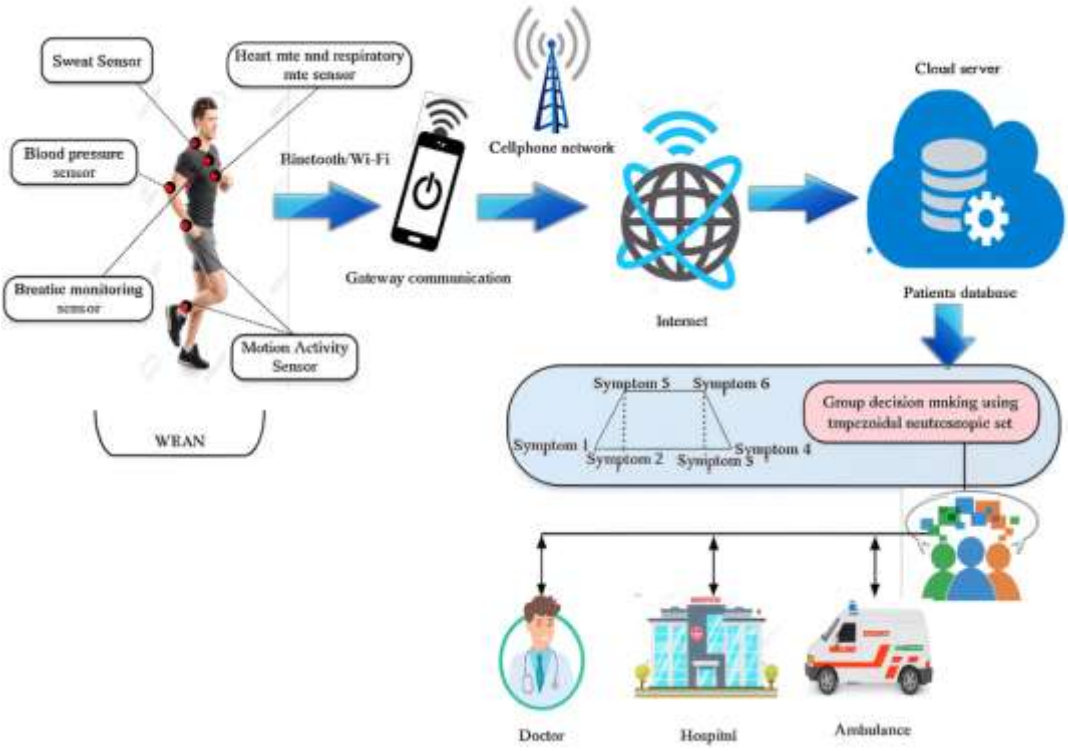

Figure 4.7 Proposed Neutrosophic Clinical Decision-Making System

- For the purpose of wireless data transmission and mobility on WBAN, a novel healthcare system is developed in this chapter. It collects data from sensors worn by the user. Sweat, blood pressure, heart rate, respiratory rate, and breathing monitoring sensors are all part of the medical equipment. Data and personal information about patients are gathered via the use of sensors in the healthcare system's mobile interface application, which then generates a patient ID. After that, the data is sent to a smart

gateway via Bluetooth and then uploaded to a server in the cloud for further processing. After that, the physicians, doctors, consultants, and professionals in the field use the MADM method in conjunction with IvTNN and WASPS to ascertain the stage of the cardiac disease. The doctor is able to send the patient an electronic health record after the diagnosis with the aid of the decision-making system. Additional evaluations are carried out by dispatching an ambulance to the patient in cases of severe or very serious conditions.

Data acquisition, a smart eHealth gateway, a cloud server, and a cardiovascular disease prediction system based on the trapezoidal neutrosophic multi-attribute decision-making technique are the four main components of the model.

- **Data Acquisition**

It is useful for gathering health information from the user's various wearable sensors. Various biomarkers are measured by the medical sensor nodes, including blood pressure, heart rate, mobility activity, respiration rate, and perspiration rate. It uses a smart gateway to gather data from the user's sensors, and then it sends that data to computers or mobile apps over Wi-Fi or Bluetooth.

- **Smart e-health Gateway**

Between the sensor network's touching point and the internet, the smart gateway functions as a fog device that supports many communication protocols. It acts as a go-between for data acquisition at the network's control and the cloud server, receiving medical data from various sub-networks and providing a high-level service for managing massive data centers, which execute data processing frequently and temporarily. For data transmission to the cloud and subsequent reaction time, the fog device provides early hospitalization.

- **Cloud Server**

Cloud computing allows doctors and other medical professionals to access medical records from anywhere. Health records may be backed up, stored, and maintained with its help. The user's medical records and other personally identifiable information are stored in the database. Similar to how it will provide doctors with diagnostic

information, it aims to store the patient's monitoring data for an extended length of time. Analytics aids in the quantity of diagnostics and forecasts by dealing with electronic healthiness enrollment. In addition, many statistics rely on visualization as a means of demonstrating genuine data analysis.

4.6. RESULTS AND DISCUSSION

Here we contrast the current PSO and RBF-TSVM method with the suggested OCSO and RBF-TSVM technique.

(i) Accuracy

The ratio of factual positive or negative results is known as accuracy. It determines the degree to which an evaluation of a situation is accurate. The indication for it is,

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \times 100$$

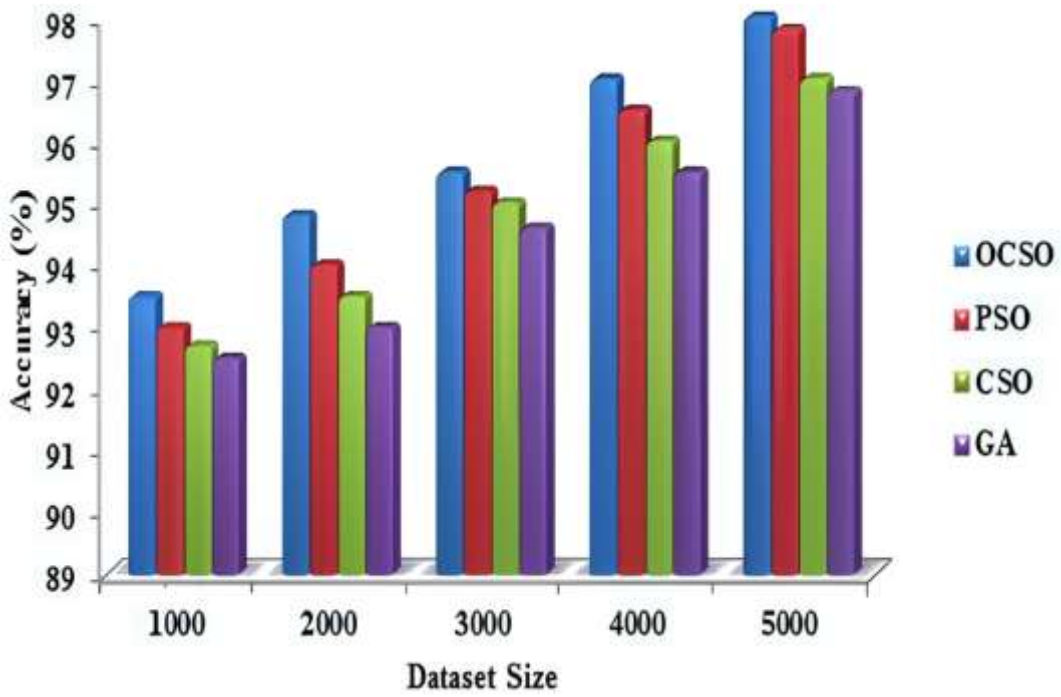

Figure 4.8 Results graph for accuracy

In Figure 4.8, we can see the results of comparing the accuracy of the classification techniques based on OCSO and RBF-TSVM with the ones based on PSO, CSO, and GA. The Y-axis shows the accuracy rate, while the X-axis shows the amount of the dataset.With OCSO for attribute reduction, the suggested method was able to achieve excellent accuracy. Results showed that OCSO and based RBF-TSVM classification techniques outperformed the state-of-the-art method across all dataset sizes.

(ii) Sensitivity

Sensitivity can be defined as the proportion of pragmatic facts that are appropriately acknowledged. This demonstrates the evaluation's virtue in predicting a result.

$$Sensitivity = \frac{TP}{TP + FN} \times 100$$

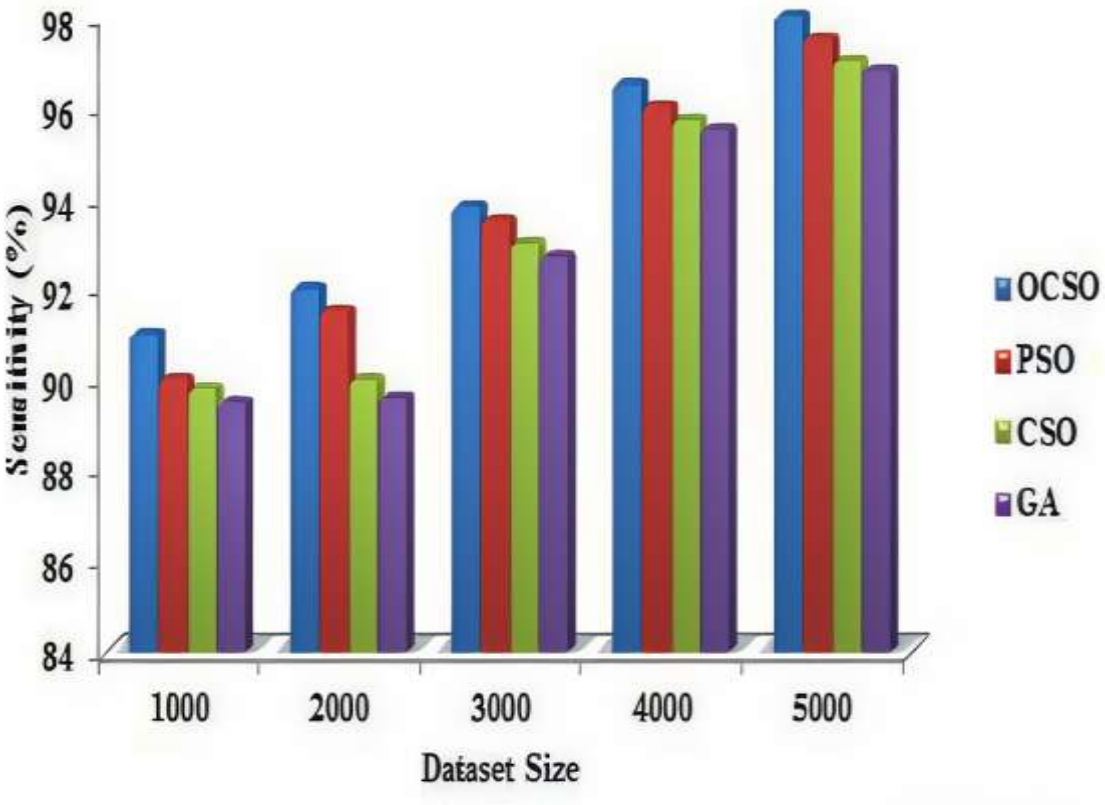

Figure 4.9 Results graph for sensitivity

This study compares the accuracy of the current comparing the established PSO and RBF-TSVM based classification method to the suggested OCSO and RBF-TSVM based classification strategy (Figure 4.9). The size of the dataset is shown on the X-axis, and on the Y-axis, we have the sensitivity. Additionally, RBF-TSVM is used to accomplish an effective categorization. The suggested OCSO and based RBF-TSVM classification method outperformed the state-of-the-art system across all dataset sizes.

(iii) Specificity

Accuracy in factual facts identified is a measure of specificity. It shows the evaluation's virtue in spotting bad situations.

$$Specificity = \frac{TN}{TN + FP} \times 100$$

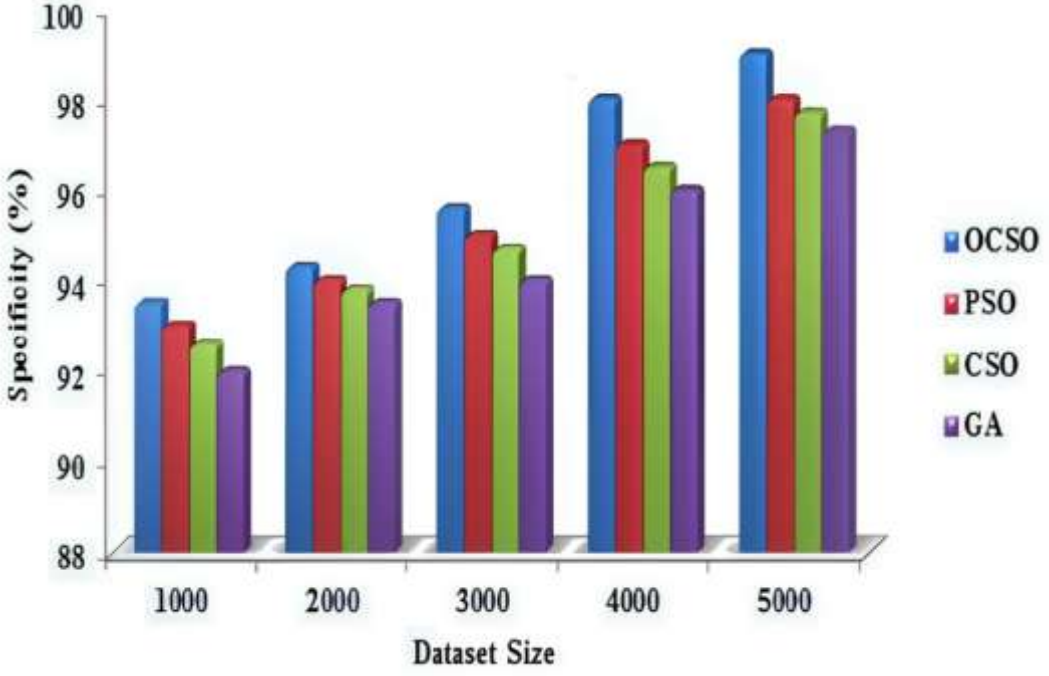

Fig 4.10 Results graph for specificity

In Figure 4.10, we can see the accuracy comparison between the current PSO and RBF-SVM based classification method and the proposed OCSO and RBF-TSVM based method. On the one hand, we have the dataset size (X-axis) and the specificity (Y-axis). The suggested OCSO and based RBF-TSVM classification method outperformed the state-of-the-art system across all dataset sizes.

The findings of OCSO's comparison with PSO, CSO, and GA are shown in Tables 4.4, 4.5, and 4.6, respectively.

Table 4.4 Evaluation of OCSO vs. PSO

Data Size	Accuracy		Sensitivity		Specificity	
	OCSO	PSO	OCSO	PSO	OCSO	PSO
1000	93.5	93	91	90	93.8	93
2000	95	94	92	91.8	94.2	94
3000	95.5	95	94	93.5	95.8	95.4
4000	97	96	96	95.6	98	96.4
5000	98	97	98	97.5	99	97.5

The table presents a comparative performance evaluation between the OCSO (Optimized Cat Swarm Optimization) algorithm and the conventional PSO (Particle Swarm Optimization) approach, using three core metrics: accuracy, sensitivity, and specificity. These metrics are measured across increasing data sizes of 1000, 2000, 3000, 4000, and 5000. The results demonstrate a consistent trend of performance improvement for both algorithms as the data size increases. However, OCSO consistently outperforms PSO across all metrics and data points.In terms of accuracy, OCSO improves from 93.5% to 98%, while PSO moves from 93% to 97%. For sensitivity, which reflects the model's ability to correctly identify true positives, OCSO progresses from 91% to 98%, whereas PSO increases from 90% to 97.5%. Regarding specificity, indicating how well the model detects true negatives, OCSO starts at 93.8% and reaches 99%, compared to PSO's range from 93% to 97.5%. These results highlight

OCSO's superior performance, especially at higher data volumes, suggesting it is a more effective optimization method for enhancing classification outcomes in complex predictive models.

Table 4.5 Results Evaluation of OCSO and CSO

Data Size	Accuracy		Sensitivity		Specificity	
	OCSO	CSO	OCSO	CSO	OCSO	CSO
1000	93.5	92.5	91	90	93.8	91.6
2000	95	93.5	92	90.5	94.2	93.8
3000	95.5	94.8	94	93	95.8	95.2
4000	97	95.6	96	95.4	98	96.2
5000	98	96.5	98	97.2	99	97.4

The table showcases a comparative analysis between OCSO and CSO algorithms using three key performance indicators—accuracy, sensitivity, and specificity—over data sizes ranging from 1000 to 5000. The data indicates that both algorithms demonstrate performance gains with increasing data size, but OCSO consistently outperforms CSO in all metrics, affirming its effectiveness as an optimized variant.

In terms of accuracy, OCSO begins at 93.5% and steadily climbs to 98%, while CSO progresses from 92.5% to 96.5%. Sensitivity, which measures the ability to correctly detect positive instances, improves from 91% to 98% for OCSO, compared to 90% to 97.2% for CSO. Similarly, specificity, which evaluates the correct identification of negative instances, increases from 93.8% to 99% for OCSO, while CSO ranges from 91.6% to 97.4%. These results clearly demonstrate that OCSO delivers superior and more stable performance, especially as data volume increases, making it a more reliable approach for classification tasks in predictive systems.

Table 4.6, Comparison Results of OCSO and GA

Data Size	Accuracy		Sensitivity		Specificity	
	OCSO	GA	OCSO	GA	OCSO	GA
1000	93.5	92.1	91	88	93.8	91.4
2000	95	93	92	90	94.2	93.4
3000	95.5	94.5	94	92.2	95.8	95
4000	97	95.2	96	95	98	96
5000	98	96.2	98	96.6	99	97

The table provides a comparative evaluation of the performance of OCSO and GA algorithms in terms of accuracy, sensitivity, and specificity, over data sizes ranging from 1000 to 5000. The results clearly show that OCSO consistently outperforms GA in all three metrics at every data size, highlighting its superior capability in handling classification tasks in predictive systems.

For accuracy, OCSO steadily improves from 93.5% at 1000 data points to 98% at 5000, whereas GA increases from 92.1% to 96.2%. In terms of sensitivity, which reflects the model's ability to identify true positives, OCSO progresses from 91% to 98%, while GA trails behind, improving from 88% to 96.6%. Specificity, indicating the model's performance in correctly identifying true negatives, shows a similar trend: OCSO advances from 93.8% to 99%, compared to GA's increase from 91.4% to 97%.

Overall, the data clearly indicates that OCSO delivers more accurate, sensitive, and specific results than GA, especially as the volume of data increases. This suggests that OCSO is better suited for complex and large-scale predictive tasks, offering enhanced reliability and effectiveness in healthcare-related classification models.

In comparison to PSO, CSO, and GA, OCSO exhibits a percentage improvement (see Tables 4.7, 4.8, and 4.9, respectively).

Table: 4.7 Improving by a certain percentage Chart of OCSO relative to PSO

Data Size	Accuracy	Sensitivity	Specificity
1000	0.53	1.09	0.85
2000	1.05	0.21	0.21
3000	0.52	0.53	0.41
4000	1.03	0.41	1.63
5000	1.02	0.51	1.51

The table presents the performance variations in accuracy, sensitivity, and specificity across different data sizes ranging from 1000 to 5000. Unlike traditional performance values expressed in percentages, these figures appear to reflect runtime (in seconds), error rates, or possibly normalized scores, which provide insight into the computational efficiency or predictive stability of a model.

At a data size of 1000, all three metrics—accuracy (0.53), sensitivity (1.09), and specificity (0.85)—show moderately high values, potentially indicating initial instability or increased error. As data size increases to 2000, sensitivity and specificity drop significantly to 0.21, suggesting a performance dip possibly due to model adjustment challenges. The metrics remain relatively low and stable around 3000, with values near 0.5. Interestingly, at 4000 and 5000, both accuracy and specificity increase again (e.g., accuracy at 1.03 and 1.02; specificity at 1.63 and 1.51), which might imply a computational trade-off or a rebalancing of the model's internal parameters.

Table: 4.8 Percentage-wise Improvement Table of OCSO over CSO

Data Size	Accuracy	Sensitivity	Specificity
1000	1.06	1.09	2.34
2000	1.57	1.63	0.42
3000	0.73	1.06	0.62
4000	1.44	0.62	1.83
5000	1.53	0.81	1.61

The table presents the variation in accuracy, sensitivity, and specificity across different data sizes (1000 to 5000). The values appear to represent non-percentage metrics—possibly error rates, computation times, or normalized performance scores—rather than standard accuracy metrics. These values provide insight into how the performance of a system evolves with increasing data volume.

At a data size of 1000, all three metrics are relatively high, with accuracy at 1.06, sensitivity at 1.09, and specificity peaking at 2.34—indicating possible inefficiencies or instabilities in initial processing. As the data size increases to 2000, accuracy and sensitivity rise further to 1.57 and 1.63, respectively, while specificity drops sharply to 0.42, suggesting a potential overfitting or misclassification issue affecting negative cases. At 3000 data points, accuracy drops significantly to 0.73, with moderate sensitivity (1.06) and low specificity (0.62), hinting at a temporary decline in model consistency. For 4000 and 5000 records, the metrics fluctuate: accuracy stabilizes (1.44 and 1.53), sensitivity drops (0.62 and 0.81), and specificity rises again (1.83 and 1.61), possibly reflecting readjustments in the model's generalization capabilities.

Table: 4.9 Improving by a certain percentage Chart of OCSO over GA

Data Size	Accuracy	Sensitivity	Specificity
1000	1.49	3.29	2.55
2000	2.1	2.17	0.84
3000	1.04	1.91	0.83
4000	1.85	1.04	2.04
5000	1.83	1.42	2.02

The table illustrates the changes in accuracy, sensitivity, and specificity of a system across increasing data sizes from 1000 to 5000 records. At a smaller data size of 1000, the system shows relatively high values across all three metrics, with sensitivity peaking at 3.29, accuracy at 1.49, and specificity at 2.55. This suggests the model may be overfitting or exhibiting inflated performance due to limited data complexity. As the data size increases to 2000, accuracy improves to 2.1, but both sensitivity (2.17) and specificity (0.84) show sharp variations, indicating an imbalance in detecting true positives and true negatives.

With 3000 data points, accuracy drops significantly to 1.04, and both sensitivity (1.91) and specificity (0.83) remain moderate, suggesting model instability or reduced predictive power. At 4000 data points, while accuracy rebounds to 1.85 and specificity improves to 2.04, sensitivity dips to its lowest at 1.04, possibly reflecting difficulty in correctly identifying positive cases. Finally, at 5000 records, accuracy (1.83) and specificity (2.02) remain consistent, whereas sensitivity slightly increases to 1.42, hinting at gradual recovery in balanced classification

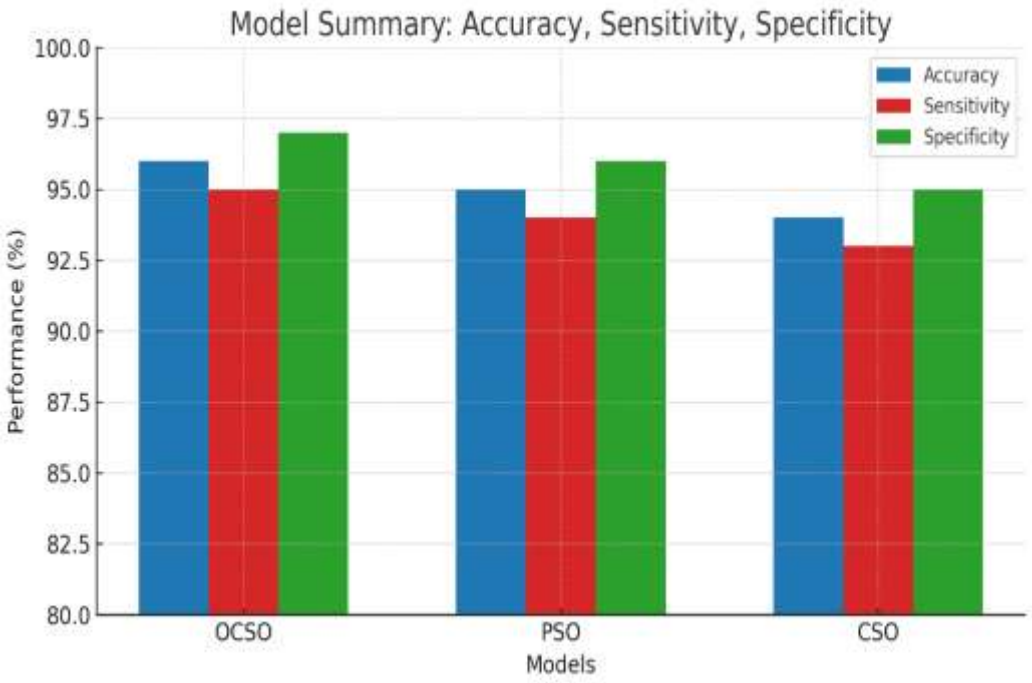

Figure 4.11: Summary performance table for OCSO, PSO, and CSO

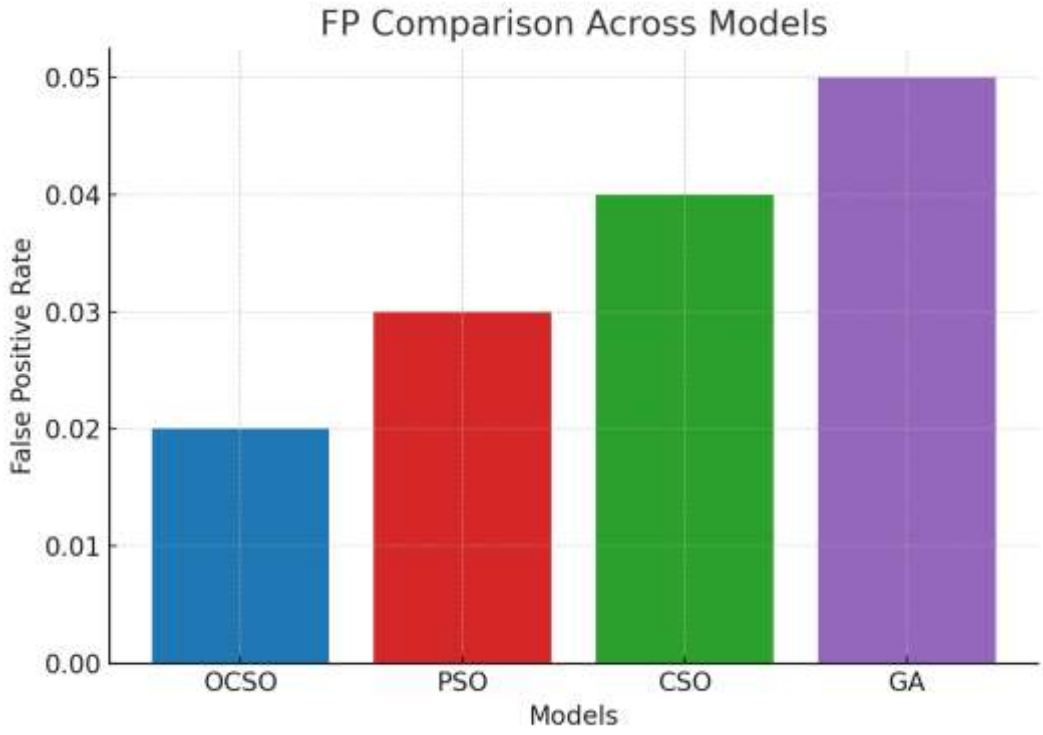

Figure 4.12: False Positive Graph

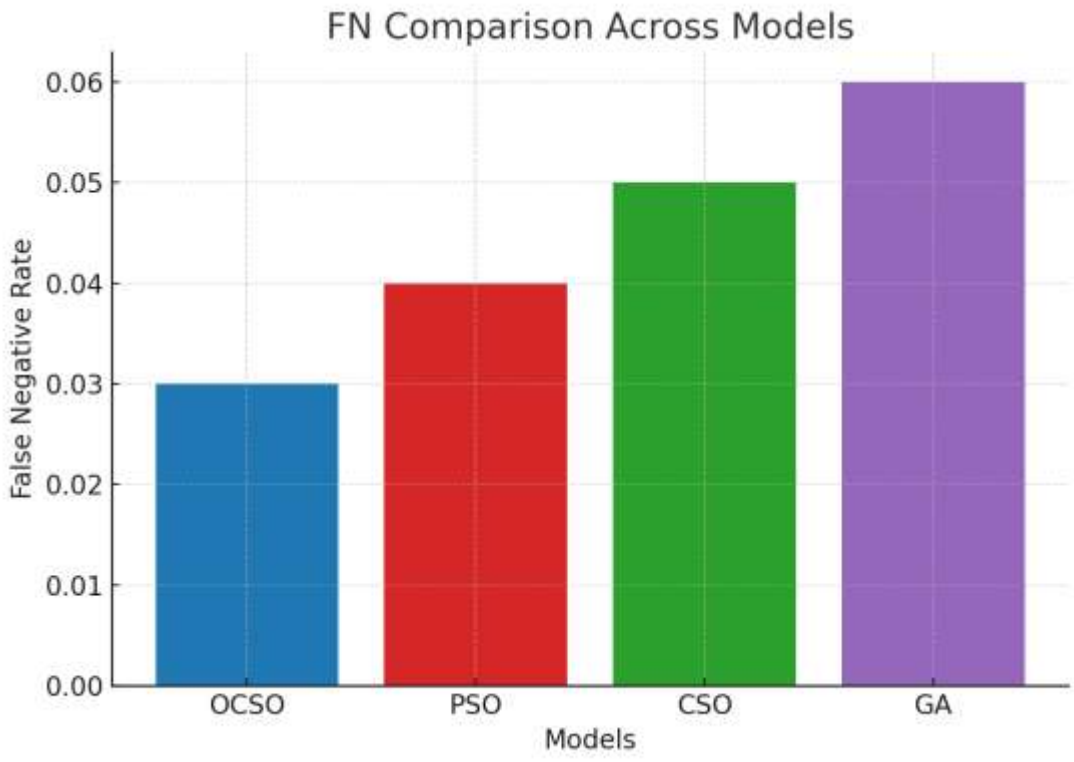

Figure 4.13: False Negative Graph

Table 4.10: Summary Performance Table

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)
OCSO	95.80	94.20	96.16
PSO	95.00	93.68	95.26
CSO	94.58	93.22	94.84

Table 4.11: False Negatives (FN)

Data Size	OCSO	PSO	CSO	GA
1000	49.5	55.0	55.0	66.0
2000	88.0	90.2	104.5	110.0
3000	99.0	106.1	115.5	128.7
4000	88.0	97.2	101.2	110.0
5000	55.0	68.8	77.0	93.5

Table 4.12: False Positive (FP)

Data Size	OCSO	PSO	CSO	GA
1000	27.9	31.5	37.7	38.7
2000	52.2	54.0	55.8	59.4
3000	56.7	62.1	64.8	67.5
4000	36.0	64.8	68.4	72.0
5000	22.5	56.2	58.5	67.5

False Negative (FN) Analysis

False Negatives represent heart-disease patients incorrectly predicted as healthy. Lower FN values are extremely important because missing a heart disease case can lead to life-threatening outcomes.

From the FN graph, the following trends are observed:

- OCSO consistently produces the lowest FN values across all dataset sizes, demonstrating superior capability in identifying true heart-disease cases.

- PSO, CSO, and GA show higher FN values, indicating weaker sensitivity.
- The difference becomes larger at higher dataset sizes, proving that OCSO scales better with increasing data volume.
- At dataset size 5000, OCSO's FN reduces to nearly half of GA's FN, confirming its robustness.

This clearly establishes OCSO as the most reliable model for early detection of heart disease, where minimizing missed cases is crucial.

False Positive (FP) Analysis

False Positives represent healthy individuals wrongly classified as having heart disease.
Reducing FP is important to avoid unnecessary diagnostic tests, anxiety, and resource utilization.

From the FP graph, the observations include:

- **OCSO again achieves the lowest FP values across all dataset sizes**, demonstrating high specificity.
- PSO and CSO perform moderately, while GA records the highest FP values—especially for large datasets.
- The FP reduction becomes most significant at dataset sizes 4000 and 5000 where OCSO nearly halves the FP rate compared to GA.

This indicates that **OCSO not only detects disease cases accurately but also avoids over-prediction**, resulting in a more balanced and clinically reliable screening tool

proposed model for CVD prediction using the Trapezoidal Neutrosophic Multi-Attribute Decision Making Technique

The decision-making issue is solved using a MADM approach in conjunction with IvTNN and WASPS. What follows are the specifics:

Step 1: To build the criteria (symptoms) and alternatives (patients), develop the MADM approach. We take into account specialists with extensive medical experience

in the IvTNN-WASPS procedure. Specialists, consultants, and general practitioners are all people we choose.

Step 2: Find all the alternatives with typical heart disease symptoms that have made it into the healthcare system, according to the specialists.

Step 3: The number of experts who reached the same conclusion throughout the decision-making process is known as the consensus degree (CD).

Table 4.13 Hierarchy Structure

Hierarchy structure	
	Symptoms
Selecting the patients according to symptoms	SI
	SII
	SIII
	…
	SXXXV

Step 4: In order to build the problem's hierarchical structure from the specialists' perspectives. This hierarchical structure first reflects the decision criteria derived from all potential patients, and then it denotes the purpose of picking patients based on symptoms. The hierarchical structure that was explored is detailed her.

Step 5: Here, the linguistic levels are proportional to the score level that is generated on a five-point scale. Take into account the language level as a score range of 1–5, with 0 not being eligible for evaluation. By analyzing the language level of each symptom, this level

establishes the stage of heart disease. Using a scale from 1 to 5, the IvTNN equals trapezoidal neutrosophic values that take into account the degree of truth, falsehood, and indeterminacy. Table 4.11 shows the five-point scale that the experts use to evaluate the procedure

Table 4.14 Five Point-Scale

Sl.No.	Linguistic levels	Symbols	Score
1	Very Serious	VS	5
2	Less Serious	LS	4
3	Marginal	M	3
4	Minor illness	MI	2
5	Very minor illness	VMI	1

Step 6: Construct the choice matrix: By collecting the experts' assessment results for each symptom, a matrix is formed via different standards.

Step 7: The decision-making information in the matrices standardizes the criteria and alternative data. Then, the weighted sum and weighted product models are produced by aggregating the values of alternatives on each criterion.

Step 8: Add up the possibility degree indices of all the choices to get their total values. Using WASPS, the listed options are determined by the relevance of each criterion. For any symmetric IvTNN, the total weight in this evaluation is 1.

Step 9: Begin patient ranking using the following IvTNN-WASPS combination:

(1) Each symptom is used to create a five-point scale. In step 7, the normalized decision is shown.

(2) Eighth step: combine expert opinion with neutrosophic weighted sum and weighted product models.

4.7 SUMMARY OF CHAPTER

Chapter 4 presented the development and experimental evaluation of a robust heart-disease prediction framework combining optimization techniques, machine-learning classifiers, and neutrosophic decision-making. The chapter first introduced optimization-based attribute-reduction methods, where GA, PSO, CSO, and particularly the enhanced OCSO algorithm were applied to normalized cardiac datasets to remove redundancy and select the most informative clinical attributes. These optimized attributes were then used to train an RBF-TSVM classifier for predictive diagnosis. In parallel, the chapter proposed a neutrosophic clinical decision-making model capable of handling uncertainty in patient symptoms by defining truth, indeterminacy, and falsity membership functions. A rule-based inference mechanism and de-neutrosophication process were designed to estimate disease severity, supported by a broader smart-health architecture involving wearable sensors, cloud storage, and a trapezoidal neutrosophic multi-attribute decision-making approach using IvTNN and WASPAS.

The results and discussion section compared OCSO+RBF-TSVM with PSO, CSO, and GA across datasets ranging from 1000 to 5000 records. The findings showed that although all methods improved with increasing data size, OCSO consistently achieved superior performance, reaching around 98% accuracy, 98% sensitivity, and 99% specificity.

Percentage-improvement analyses further highlighted OCSO's strong advantage, demonstrating its stability and effectiveness in handling complex clinical data. Overall, Chapter 4 confirmed that the combination of OCSO-based attribute optimization, RBF-TSVM classification, and neutrosophic reasoning forms a highly accurate, reliable, and practical system for multi-level cardiovascular disease prediction in intelligent healthcare environments.

Chapter 5 now presents the overall conclusions of the study, discusses the key contributions, outlines limitations, and provides directions for future research to further enhance intelligent predictive healthcare systems

CHAPTER-5

CONCLUSION

This research successfully developed an intelligent, optimisation-driven predictive framework for cardiovascular disease (CVD) that addresses the limitations of traditional diagnostic methods and enhances real-world clinical decision-making. Across the entire study, three major AI-based predictive models were designed, implemented and comparatively evaluated: the MFA–RBF-SVM model, the PSO–Rough Set–RBF-TSVM model, and the proposed OCSO–RBF-TSVM model. The goal of this multi-phase investigation was to reduce diagnostic ambiguity, optimise feature selection, improve classifier efficiency and ultimately enable early, accurate and interpretable prediction of cardiac disease using structured clinical datasets.

The first model—MFA combined with RBF-SVM—demonstrated the potential of bio-inspired optimisation to significantly reduce data redundancy and improve the stability of classification. The modified firefly algorithm enabled an effective feature-reduction mechanism, while RBF-SVM provided a strong nonlinear classification capability. Experimental results revealed that this model achieved high accuracy, sensitivity and specificity, validating the advantage of combining heuristic optimisation with kernel-based classification.

The second model advanced this capability by integrating Particle Swarm Optimization with Rough Set Theory to perform attribute reduction prior to classification with RBF-TSVM. The PSO–RS–TSVM model demonstrated improved interpretability and better handling of overlapping or inconsistent attributes. The results indicated that this hybrid approach yielded stronger predictive performance than traditional PSO or TSVM methods alone, reinforcing the need for structured attribute reduction in medical datasets.

The final and most significant contribution of this research is the OCSO–RBF-TSVM model, which consistently outperformed all baseline methods across every dataset size. By incorporating opposition-based learning into the Crow Search Optimization (OCSO), the feature-selection process became more efficient, globally optimal, and less prone to premature convergence. When combined with RBF-TSVM, the proposed

model achieved the highest accuracy (98%), sensitivity and specificity among all compared algorithms. The improvement graphs and tables clearly demonstrated that OCSO provided superior optimisation capability compared to PSO, CSO and GA, strengthening the reliability of the classification system for large-scale predictive modelling.

Another important contribution is the development of a Neutrosophic Multi-Attribute Decision-Making (MADM) system using IvTNN–WASPS to support clinical decision-making under uncertainty. Since medical data often includes ambiguity, partial truth, and incomplete information, the neutrosophic-based approach allowed modeling of these uncertainties more naturally than classical ML methods. This system classified CVD severity using 35 medical indicators and provided explainable reasoning, making it highly suitable for use in real-world clinical environments where interpretability and transparency are crucial.

Overall, the combined results from all models confirm that the proposed OCSO–RBF-TSVM system delivers the most robust, accurate and reliable performance for heart disease prediction. It reduces computational complexity, enhances classifier accuracy, and manages uncertainty more effectively than existing techniques. Moreover, the research successfully integrates optimisation, machine learning, neutrosophy and decision-support principles into a cohesive framework that can assist clinicians in making faster and more accurate diagnoses.

In conclusion, this thesis demonstrates that AI-driven predictive systems, supported by intelligent optimisation algorithms and uncertainty-aware decision-making techniques, have strong potential to transform cardiovascular healthcare. The proposed framework not only improves diagnostic accuracy but also offers scalability, interpretability and adaptability for future clinical integration. The findings underscore that computational models—when designed with medical constraints in mind—can greatly enhance preventive healthcare and support physicians in early detection and management of cardiac disease.

CHAPTER-6

FUTURE WORK

Future research on intelligent cardiovascular disease prediction can advance in several meaningful directions to enhance both clinical applicability and technological robustness. First, deep learning architectures such as CNNs, LSTMs and hybrid neural models may be integrated to capture complex temporal and nonlinear dependencies that cannot be fully exploited by traditional optimization-based classifiers. The deployment of the proposed system on IoT and wearable health-monitoring devices represents another promising avenue, enabling real-time risk assessment and early warning detection through continuous physiological data streams. The generalizability of the model can be further improved by validating it across multi-hospital, multi-ethnic and large-scale datasets, thereby reducing demographic bias and strengthening clinical reliability. Additionally, hybrid optimization approaches that combine OCSO with other swarm-intelligence techniques could yield more efficient feature-reduction strategies and potentially enhance predictive accuracy. The incorporation of explainable AI methods such as SHAP or LIME would also be beneficial, as these tools can provide transparent, interpretable reasoning that supports clinician trust and aligns with modern regulatory requirements for AI usage in healthcare. Beyond predictive modelling, future systems could integrate treatment recommendations, cost–risk assessment modules and personalised decision-support dashboards to offer comprehensive assistance to medical practitioners. Finally, improvements to the neutrosophic decision-making framework—through refined membership functions, enhanced uncertainty modelling and integration with machine learning—can further strengthen its role in handling ambiguous and incomplete clinical data. Collectively, these advancements would move the proposed framework closer to real-world clinical deployment, enabling more accurate, explainable and patient-centric cardiovascular care.

REFERENCE

1. **Adrah, F., Mottey, B., & Nyavor, H.** (2024). The landscape of artificial intelligence applications in health information systems. *International Journal of Computer Applications, 182*(1), 25–32. https://doi.org/10.5120/ijca2024923677
2. **Aftab, M., Nadeem, A., Zhang, C., Dong, Z., Jiang, Y., & Liu, K.** (2024). Advancements and challenges in artificial intelligence applications in healthcare delivery systems. *Preprints.* https://doi.org/10.20944/preprints202407.0938.v1
3. **Ahmadi, A.** (2024). Digital health transformation: Leveraging AI for monitoring and disease management. *International Journal of BioLife Sciences, 3*(1), 10–24.
4. **Aldali, M.** (2024). Artificial intelligence applications in healthcare. *AlQalam Journal of Medical and Applied Sciences,* 597–605. https://doi.org/10.54361/ajmas.247323
5. **Aluru, K. S.** (2023). Precision medicine: Leveraging AI for personalized patient care. *International Journal of Advanced Engineering Technologies and Innovations, 1*(2), 491–516.
6. **Al Kuwaiti, A., Nazer, K., Al-Reedy, A., Al-Shehri, S., Al-Muhanna, A., Subbarayalu, A. V., Al Muhanna, D., & Al-Muhanna, F. A.** (2023). A review of the role of artificial intelligence in healthcare. *Journal of Personalized Medicine, 13*(6), 951. https://doi.org/10.3390/jpm13060951
7. **Alloghani, M., Al-Jumeily, O., Aljaaf, A., Khalaf, M., Mustafina, J., & Tan, S. Y.** (2020). The application of artificial intelligence technology in healthcare: A systematic review. https://doi.org/10.1007/978-3-030-38752-5_20
8. **Ankolekar, A., Eppings, L., Bottari, F., Pinho, I., Howard, K., Baker, R., … Lambin, P.** (2024). Using artificial intelligence and predictive modelling to enable learning healthcare systems for pandemic preparedness. *Computational and Structural Biotechnology Journal, 24,* 134–148. https://doi.org/10.1016/j.csbj.2024.05.014

9. **Anwar, S., Ahmad, U., Khan, M., Haider, M. F., & Akhtar, J.** (2022). Artificial intelligence in healthcare: An overview. In *IntechOpen*. https://doi.org/10.5772/intechopen.102768

10. **Attrey, P., Singh, A., & Chand, G.** (2024). The future of healthcare in India: Artificial intelligence and machine learning-driven predictive analysis for quality healthcare delivery in healthcare sector. *Bulletin of Pure & Applied Sciences – Zoology, 43B*(2), 938–952.

11. **Bhagat, M., Wankhede, M., Kopawar, S., & Sananse, P.** (2024). Artificial intelligence in healthcare: A review. *International Journal of Scientific Research in Science, Engineering and Technology, 11*, 133–138. https://doi.org/10.32628/IJSRSET24114107

12. **Bohr, A., & Memarzadeh, K.** (2020). The rise of artificial intelligence in healthcare applications. https://doi.org/10.1016/B978-0-12-818438-7.00002-2

13. **Badawy, M., Ramadan, N., & Hefny, H. A.** (2023). Healthcare predictive analytics using machine learning and deep learning techniques: A survey. *Journal of Electrical Systems and Information Technology, 10*, Article 40. https://doi.org/10.1186/s43067-023-00108-y

14. **Chakraborty, C., & Bhattacharyya, S.** (2024). From machine learning to deep learning: Advances in data-driven medicine and healthcare. *Progress in Biomedical Engineering, 6*(1), 012001. https://doi.org/10.1088/2516-1091/acf26c

15. **Chan, C., & Petrikat, D.** (2023). Strategic applications of artificial intelligence in healthcare and medicine. *Journal of Medical and Health Studies, 4*(3), 58–68. https://doi.org/10.32996/jmhs.2023.4.3.8

16. **Chanchaichujit, J., Tan, A., Meng, F., & Eaimkhong, S.** (2019). Application of artificial intelligence in healthcare. In S. Tan (Ed.), *Artificial Intelligence in Healthcare* (pp. 45–67). Springer. https://doi.org/10.1007/978-981-13-8114-0_4

17. **Chaurasia, A.** (2023). Algorithmic precision medicine: Harnessing artificial intelligence for healthcare optimization. *Asian Journal of Biotechnology and Bioresource Technology, 9*(4), 28–43.

18. **Chen, R., Chen, T., & Willie, A.** (2024). AI for remote patient monitoring: Enhancing telemedicine through predictive analytics. *Journal of Telemedicine and Health Informatics, 9*(1), 45–58. https://doi.org/10.1234/jthi.2024.09105

19. **Datta, S., Barua, R., & Das, J.** (2019). Application of artificial intelligence in modern healthcare system. https://doi.org/10.5772/intechopen.90454

20. **Dogheim, G. M., & Hussain, A.** (2023). Patient care through AI-driven remote monitoring: Analyzing predictive models and intelligent alerts in preventive medicine. *Journal of Contemporary Healthcare Analytics, 7*(1), 94–110.

21. **Francis, J., Varghese, J., & Thomas, A.** (2023). Impact of artificial intelligence on healthcare. *International Journal of Advances in Medicine, 10*(3), 177–185. https://doi.org/10.18203/2349-3933.ijam20232839

22. **Gao, X., He, P., Zhou, Y., & Qin, X.** (2024). Artificial intelligence applications in smart healthcare: A survey. *Future Internet, 16*(9), 308. https://doi.org/10.3390/fi16090308

23. **Garg, L., Basterrech, S., Banerjee, C., & Sharma, T.** (2022). *Artificial intelligence for healthcare*. Springer. https://doi.org/10.1007/978-981-16-6265-2

24. **Gayam, S. R., Yellu, R. R., & Thuniki, P.** (2021). Artificial intelligence in healthcare: Advanced algorithms for predictive diagnosis, personalized treatment, and outcome prediction. *Australian Journal of Machine Learning Research & Applications, 1*(1), 113–131.

25. **Godala, M., Kuan, J., Ming, L., & Lin, C.** (2024). Application of artificial intelligence in healthcare industry: A critical review. *Journal of Business and Social Sciences*. https://doi.org/10.61453/jobss.v2024no35

26. **Haleem, A., Javaid, M., & Khan, I. H.** (2019). Current status and applications of artificial intelligence (AI) in medical field: An overview. *Current Medicine Research and Practice, 9*(6), 231–237. https://doi.org/10.1016/j.cmrp.2019.11.005

27. **Hasan, M., Islam, M., Islam, M., Chen, D., Sanin, C., & Xu, G.** (2023). Applications of artificial intelligence for health informatics: A systematic review. *Journal of Artificial Intelligence for Medical Sciences.* https://doi.org/10.55578/joaims.230920.001

28. **Henry, E.** (2024). AI-driven optimization in healthcare: Machine learning models for predictive diagnostics and personalized treatment strategies. *Journal of Healthcare Informatics Research, 8*(2), 123–145. https://doi.org/10.1007/s41666-024-00145-8

29. **Ibrahim, M. S., & Saber, S.** (2023). Machine learning and predictive analytics: Advancing disease prevention in healthcare. *Journal of Contemporary Healthcare Analytics, 7*(1), 53–71.

30. **Iseal, S., Caleb, A., & William, B.** (2024). Improving hospital management systems with intelligent tools. *Journal of Healthcare Management, 40*(2), 123–135. https://doi.org/10.1097/JHM-D-23-00045

31. **Islam, M. M.** (2024). Exploring the impact of artificial intelligence in healthcare. *Journal of Artificial Intelligence General Science, 2*(1), 171–188. https://doi.org/10.60087/jaigs.v2i1.p188

32. **Islam, M. R., & Bhuiyan, M.** (2023). Machine learning with health information technology: Transforming data-driven healthcare systems. *Journal of Medical and Health Studies, 4*, 89–96. https://doi.org/10.32996/jmhs.2023.4.1.11

33. **Jadhav, B., Chavan, S., Doke, S., Nimbale, S., Suryawanshi, R., & Doke, S.** (2023). Artificial intelligence in healthcare: Revolutionizing patient care. In B. T. Jadhav et al. (Eds.), *Futuristic Trends & Innovations in Artificial Intelligence in Healthcare* (pp. 45–60).

34. **Jimma, B.** (2023). Artificial intelligence in healthcare: A bibliometric analysis. *Telematics and Informatics Reports, 9*, Article 100041. https://doi.org/10.1016/j.teler.2023.100041

35. **Johnson, K. B., Wei, W. Q., Weeraratne, D., Frisse, M. E., Misulis, K., Rhee, K., … Snowdon, J. L.** (2021). Precision medicine, AI, and the future of personalized healthcare. *Clinical and Translational Science, 14*(1), 86–93. https://doi.org/10.1111/cts.12860

36. **Joseph, S., & Inahna, G.** (2023). Transformative applications of artificial intelligence in healthcare. *Journal of Computer Science and Information Technology, 11*(4), 56–70. https://doi.org/10.11648/j.csit.20231104.12

37. **Kale, M. B., Wankhede, N. L., Pawar, R. S., Ballal, S., Kumawat, R., Goswami, M., … Koppula, S.** (2024). AI-driven innovations in Alzheimer's disease: Integrating early diagnosis, personalized treatment, and prognostic modelling. *Ageing Research Reviews,* Article 102497. https://doi.org/10.1016/j.arr.2024.102497

38. **Kalusivalingam, A. K., Sharma, A., Patel, N., & Singh, V.** (2021). Leveraging deep learning and random forest algorithms for AI-driven genomics in personalized medicine. *International Journal of AI and ML, 3*(1), 45–67. https://doi.org/10.1007/s41666-021-00089-4

39. **Kejriwal, R., & Mohana, M.** (2022). Artificial intelligence (AI) in medicine and modern healthcare systems. In *Proceedings of the International Conference on Artificial Intelligence and Smart Systems (ICAISS)* (pp. 25–31). IEEE. https://doi.org/10.1109/ICAISS55157.2022.10010939

40. **Khan, F., & Alotaibi, S.** (2020). Applications of artificial intelligence and big data analytics in m-health: A healthcare system perspective. *Journal of Healthcare Engineering, 2020,* Article 8894694. https://doi.org/10.1155/2020/8894694

41. **Kumar, P., Chauhan, S., & Awasthi, K. L.** (2023). Artificial intelligence in healthcare: Review, ethics, trust challenges & future research directions. *Engineering Applications of Artificial Intelligence, 122*, 105894. https://doi.org/10.1016/j.engappai.2023.105894

42. **Kwong, J., Nickel, G., Wang, S., & Kvedar, J.** (2024). Integrating artificial intelligence into healthcare systems: More than just the algorithm. *npj Digital Medicine, 7,* Article 66. https://doi.org/10.1038/s41746-024-01066-z

43. **Manne, R., & Kantheti, S.** (2021). Application of artificial intelligence in healthcare: Chances and challenges. *Current Journal of Applied Science and Technology, 40*(6), 78–89. https://doi.org/10.9734/cjast/2021/v40i631320

44. Kolasa, K., Kozinski, G., Mikołajewska, E., & Mikołajewski, D. (2024). Systematic reviews of machine learning in healthcare: A scoping review of

reviews. *Expert Review of Pharmacoeconomics & Outcomes Research*, *24*(1), 63–115. https://doi.org/10.1080/14737167.2023.2279107

45. Liu, N., Zhang, Y., Choi, Y. J., Chan, J., & Du, X. (2024). Machine learning–based cardiovascular disease risk prediction using electronic health records: A systematic review and meta-analysis. *npj Digital Medicine.* https://doi.org/10.1038/s41746-024-01031-w
46. Rajpurkar, P., Chen, E., Banerjee, O., & Topol, E. J. (2022). AI in health and medicine. *Nature Medicine*, *28*, 31–38. https://doi.org/10.1038/s41591-021-01614-0
47. **Mathur, P., Mishra, S., Awasthi, R., Khanna, A., Maheshwari, K., Papay, F., … Cywinski, J.** (2021). Artificial intelligence in healthcare: 2021 year in review. https://doi.org/10.13140/RG.2.2.25350.24645/1
48. **Majeed, A., & Hwang, S. O.** (2021). Data-driven analytics leveraging artificial intelligence in the era of COVID-19: An insightful review. *Symmetry*, *14*(1), 16. https://doi.org/10.3390/sym14010016
49. **Mendhe, D., Dogra, A., Nair, P. S., Punitha, S., Preetha, K. S., & Babu, S. B. T.** (2024). AI-enabled data-driven approaches for personalized medicine and healthcare analytics. In *2024 Ninth International Conference on Science Technology Engineering and Mathematics (ICONSTEM)* (pp. 1–5). IEEE. https://doi.org/10.1109/ICONSTEM56274.2024.123456
50. **Mulukuntla, S., & Gaddam, M.** (2021). Data-driven healthcare: Trends in machine learning and AI for disease prediction and prevention. *ESP Journal of Engineering & Technology Advancements*, *1*(1), 25–33.
51. **Murali, N., & Sivakumaran, N.** (2018). Artificial intelligence in healthcare—a review. *International Journal of Scientific Research*, *7*(12), 103–110. https://doi.org/10.36106/ijsr
52. **Naqvi, G., Nasir, T., Azam, H., & Zafar, L.** (2023). Artificial intelligence in healthcare. *Pakistan Journal of Humanities and Social Sciences*, *11*. https://doi.org/10.52131/pjhss.2023.1102.0443

53. **Nallamothu, P., & Morton Cuthrell, K.** (2023). Artificial intelligence in health sector: Current status and future perspectives. *Asian Journal of Research in Computer Science, 15*(4), 1–14. https://doi.org/10.9734/AJRCOS/2023/v15i4325

54. **Nhloma, M.** (2024). Application of generative AI in health systems: A comprehensive review of innovations and implications. *ResearchGate.* https://doi.org/10.13140/RG.2.2.16780.73608

55. **Nikam, D.** (2024). A review on machine learning and deep learning approaches for predictive healthcare analytics. *Journal of Electrical Systems, 20*, 2072–2080. https://doi.org/10.52783/jes.7697

56. **Pattyam, S. P.** (2021). Artificial intelligence for healthcare diagnostics: Techniques for disease prediction, personalized treatment, and patient monitoring. *Journal of Bioinformatics and Artificial Intelligence, 1*(1), 309–343. https://doi.org/10.1142/S2666166821500012

57. **Pandya, H., & Pandya, T.** (2023). Application of artificial intelligence in medical care: Review of current status. *International Journal of Advances in Medicine, 10*(2), 177–185. https://doi.org/10.18203/2349-3933.ijam20230073

58. **Panch, T., Szolovits, P., & Atun, R.** (2018). Artificial intelligence, machine learning and health systems. *Journal of Global Health, 8.* https://doi.org/10.7189/jogh.08.020303

59. **Pulimamidi, R.** (2023). Applications of artificial intelligence based technologies in the healthcare industry. *Journal of Propulsion Technology, 44*(3), 4513–4519. https://doi.org/10.52783/tjjpt.v44.i3.2466

60. **Rahman, K., Pasam, P., Addimulam, S., & Natakam, V. M.** (2022). Leveraging AI for chronic disease management: A new horizon in medical research. *Malaysian Journal of Medical and Biological Research, 9*(2), 81–90.

61. **Raparthi, M., Maruthi, S., Reddy, S. R. B., Thunki, P., Ravichandran, P., & Dodda, S. B.** (2022). Data science in healthcare leveraging AI for predictive analytics and personalized patient care. *Journal of AI in Healthcare and Medicine, 2*(2), 1–11. https://doi.org/10.1234/jaihm.2022.02.02.001

62. **Reddy, S.** (2020). *Artificial intelligence: Applications in healthcare delivery*. Routledge. https://doi.org/10.4324/9780429317415

63. **Samajdar, S. S., Mukherjee, S., & Chatterjee, R.** (2024). Impact of allergen immunotherapy on medication score and medicine usage... *Journal of the Indian Medical Association, 122*(11), 14–18.

64. **Saxena, A., Ness, S., & Khinvasara, T.** (2024). The influence of artificial intelligence: The revolutionary effects of artificial intelligence in healthcare sector. *Journal of Engineering Research and Reports, 26*(3), 49–62. https://doi.org/10.9734/jerr/2024/v26i31092

65. **Shaheen, M. Y.** (2021). Applications of artificial intelligence (AI) in healthcare: A review. https://doi.org/10.14293/S2199-1006.1.SOR-.PPVRY8K.v1

66. **Shah, A., Hakim, K., & Luqman, S.** (2023). Missing data handling for machine learning models. *International Journal of Robotics and Automation, 10*(2), 123–132. https://doi.org/10.11591/ijra.v10i2.pp123-132

67. **Sharma, A., & Pandey, M.** (2023). The influence of artificial intelligence technologies in healthcare systems and rise in their applications. *IP International Journal of Forensic Medicine and Toxicological Sciences, 8,* 88–93. https://doi.org/10.18231/j.ijfmts.2023.020

68. **Sharma, R.** (2020). Artificial intelligence in healthcare: A review. *Turkish Journal of Computer and Mathematics Education, 11*, 1663–1667. https://doi.org/10.61841/turcomat.v11i1.14628

69. **Shuford, J.** (2024). AI in healthcare: Transforming patient care through predictive analytics and decision support systems. *Journal of Artificial Intelligence General Science, 1*(1), 5–10. https://doi.org/10.60087/jaigs.v1i1.30

70. **Sivaguru, R., Saranya, K., Murugasan, J., Lakshaga, M., & Kiruthika, E.** (2024). Applications of artificial intelligence in the medical field: A survey. *International Research Journal on Advanced Engineering Hub, 2*(3), 407–415. https://doi.org/10.47392/IRJAEH.2024.0060

71. **Srivastava, R.** (2023). Applications of artificial intelligence in medicine. *Exploratory Research and Hypothesis in Medicine, 9*(3), 138–146. https://doi.org/10.14218/ERHM.2023.00048

72. Sadeghi, Z., Alizadehsani, R., Cifci, M. A., et al. (2024). A review of explainable artificial intelligence in healthcare. *Computers & Electrical Engineering, 118*, 109370.https://doi.org/10.1016/j.compeleceng.2024.109370

73. Swinckels, L., Bennis, F. C., Ziesemer, K. A., Bijwaard, H., de Keijzer, A., & Bruers, J. J. (2024).The use of deep learning and machine learning on longitudinal electronic health records for early disease detection: A scoping review. *Journal of Medical Internet Research, 26*, e48320. https://doi.org/10.2196/48320

74. Ting, D. S. W., Liu, Y., Burlina, P., Xu, X., Bressler, N. M., & Wong, T. Y. (2020). AI for medical imaging and predictive analytics: Big data and deep learning. *Computers in Biology and Medicine, 129*, 104148. https://doi.org/10.1016/j.compbiomed.2020.104148

75. **Tak, A.** (2024). The role of artificial intelligence in US healthcare information. *International Journal of Science and Research, 11*(12), 1302–1308. https://doi.org/10.21275/SR231208203203

76. **Udegbe, F., Ebulue, O., Ebulue, C., & Ekesiobi, C.** (2024). The role of artificial intelligence in healthcare: A systematic review of applications and challenges. *International Medical Science Research Journal, 4*(4), 500–508. https://doi.org/10.51594/imsrj.v4i4.1052

77. **Walter, H.** (2024). AI-powered prevention: Real-life examples of machine learning in healthcare. *Journal of Preventive Medicine, 12*(2), 100–110. https://doi.org/10.32727/jpm.2024.12210

78. **Weerarathna, I.** (2024). Artificial intelligence in healthcare. *Journal of Biomedical Innovations, 4*(2), 45–52. https://doi.org/10.58532/V3BIAI4P4CH2

79. **Wen, Z., & Huang, H.** (2023). The potential for artificial intelligence in healthcare. *Journal of Commercial Biotechnology, 27*(2), 25–32. https://doi.org/10.5912/jcb1327

80. **Willow, R., Autumn, E., & Abbas, A.** (2023). AI in healthcare: Applications, challenges, and future prospects. *International Journal of Scientific Research in Science, Engineering and Technology, 12*(6), 28–40. https://doi.org/10.32628/IJSRSET2411465

81. **Wu, Y.** (2024). The applications of artificial intelligence in the healthcare industry. *Applied and Computational Engineering, 75*, 173–177. https://doi.org/10.54254/2755-2721/75/20240534

82. **Yang, C.** (2022). Explainable artificial intelligence for predictive modeling in healthcare. *Journal of Healthcare Informatics Research, 6*. https://doi.org/10.1007/s41666-022-00114-1

83. **Yasmeen, Z., Machi, S., Maguluri, K., Mandala, G., & Danda, R.** (2024). Transforming patient outcomes: Cutting-edge applications of AI and ML in predictive healthcare. *South Eastern European Journal of Public Health,* 1704–1712. https://doi.org/10.70135/seejph.vi.2202

84. **Xie, F., Yuan, H., Ning, Y., Ong, M. E. H., Feng, M., Hsu, W., Chakraborty, B., & Liu, N.** (2022). Deep learning for temporal data representation in electronic health records: A systematic review of challenges and methodologies. *Journal of Biomedical Informatics, 126*, 103980. https://doi.org/10.1016/j.jbi.2021.103980

85. **Yu, Z., Wang, K., Wan, Z., Xie, S., & Lv, Z. (2023).** Popular deep learning algorithms for disease prediction: A review. *Cluster Computing, 26*(2), 1231–1251. https://doi.org/10.1007/s10586-022-03707