IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Mathematical Modelling of Engineering Problems
Vol. 13, No. 3, March, 2026, pp. 501-512
Journal homepage: http://iieta.org/journals/mmep

A Dual-Memory Cognitive Algorithm for Heart Disease Diagnosis: An Unconscious Mind-Inspired Model with Pattern Consolidation and k-Nearest Neighbor Retrieval

Ahmed Al-Qazzaz[1], Ali Qutub[2], Abrar Khaled Shukri[1], Mohammed Safar[2*], Mohammed N. Qasim[1]

[1] Artificial Intelligence Techniques Engineering Department, Technical Engineering College for Computer and AI–Kirkuk, Northern Technical University, Kirkuk 36001, Iraq
[2] Information Techniques and Computer Networks Engineering Department, Technical Engineering College for Computer and AI–Kirkuk, Northern Technical University, Kirkuk 36001, Iraq

Corresponding Author Email: mohammed.sefer@ntu.edu.iq

https://doi.org/10.18280/mmep.130305

Received: 16 December 2025
Revised: 10 March 2026
Accepted: 19 March 2026
Available online: 10 April 2026

Keywords:
heart disease diagnosis, cognitive classifier, dual-memory architecture, k-nearest neighbors, interpretable machine learning, clinical decision support

ABSTRACT

This study develops and optimizes the Unconscious Mind-Inspired Algorithm (UMIA), a cognitively inspired classifier with a dual-memory architecture, for heart disease diagnosis using the UCI Cleveland Heart Disease dataset which content 303 records where 297 are complete cases after removal of missing values the input preprocessing was performed with StandardScaler to normalize the continuous clinical variables before model training whereas the algorithm's internal idle-state processing used a separate 60% sigmoid transformation and 40% z-score blending to refine memory representations during the pattern consolidation. The method integrates novelty-based pattern filtering, Hebbian association learning, and k-nearest neighbor k-NN retrieval within a transparent memory-based reasoning framework. The dataset was divided into 237 training cases and 60 test cases using a stratified 80:20 split, and model selection was carried out on the training data using a validation subset. Performance was compared with Random Forest, Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), and Logistic Regression using accuracy, precision, recall, F1-score, and specificity. Statistical significance was evaluated using the Friedman test and paired t-tests. The optimized UMIA achieved 86.67% accuracy on the test set, matching Random Forest and exceeding SVM 85.00%, XGBoost 85.00%, Logistic Regression 83.33%, and LightGBM 80.00%. It also achieved 87.22% precision, 78.57% recall, 86.55% F1-score, and 93.75% specificity. Bayesian optimization improved performance by 18.34 percentage points in absolute accuracy over the baseline configuration, corresponding to a 26.84% relative improvement. These results show that UMIA provides competitive diagnostic performance while improving interpretability through dual-memory inspection, similarity-based case retrieval, and frequency-weighted pattern importance, which are valuable properties for clinical decision support.

1. INTRODUCTION

Cardiovascular disease (CVD) remains the biggest challenge facing modern healthcare worldwide, accounting for almost 17.9 million deaths per year and about 31% of total deaths worldwide [1]. The enormous death toll caused by CVD highlights the need for early diagnosis and accuracy as an intervention strategy for improving patient outcomes and halting the burden associated with CVD worldwide [2]. Complexity associated with the diagnosis and detection of CVD arises from its multi-component characteristics, which include interactions among genetic factors, lifestyles, and various components that must be analyzed at once within an objective diagnostic scope [3]. The conventional method associated with CVD diagnosis depends on cardiologists who must analyze historical and physical perceptions, electrocardiographic images, stress tests, and lab analysis for an objective diagnosis and interpretation [4]. Notably, this method requires human expertise and intervention and is not only expensive and cumbersome but also predisposed to various factors associated with observation bias and error rates within environments with limited access to cardiac facilities and expertise [5].

Despite extensive research on machine learning-based heart disease prediction using the Cleveland dataset, three technical gaps were found. First, ensemble methods like Random Forest and Extreme Gradient Boosting (XGBoost) achieve 85–88% accuracy but provide no transparent reasoning that interpretable models, such as Decision Trees and Logistic Regression, sacrifice 3–5 percentage points in accuracy unacceptable when both performance and transparency are required [6]. Secondly, the standard k-nearest neighbors (k-NNs) lack confidence quantification, pattern importance weighting, and memory consolidation mechanisms [7], and

lastly, the Cleveland dataset presents specific challenges as moderate size n = 297, class imbalance 54.5% positive, mixed feature types 5 continuous, 8 categorical, which require methods that address all three simultaneously.

The recent developments in machine learning and artificial intelligence bring about a paradigm shift for medical diagnosis with unprecedented opportunities for medical online screening, computer-aided clinical monitoring, pattern recognition, and enhancing medical diagnostic capabilities [8]. During the previous two decades, researchers have explored different machine learning algorithms for CVD prediction with promising outcomes that have successfully matched or even outperformed medical expert capabilities on multiple instances [9]. Methods using Support Vector Machines (SVMs) have been explored for finding optimal classification margins within high-dimensional feature space for medical problems [10]. Ensemble learning, including methods based on Random Forests and Gradient Boosting, has shown successful performance capabilities with multiple instances of weak learning mechanism aggregation [11]. Deep learning solutions ranging from Convolutional Neural Network (CNN) on medical imaging problems and Recurrent Neural Network (RNN) on medical physiological signals have shown successful implementation capabilities on specific medical diagnostic problems [12]. Despite accelerated advances in multiple aspects of medical diagnosis machine learning problems, successful implementation and adoption on large medical diagnostic problems have yet to be met due to multiple fundamental challenges beyond predictive accuracy capabilities [13].

"Black box" techniques are also among the biggest challenges to implementation because these state-of-the-art machine learning models are highly accurate but opaque about their predictions and reasoning processes [14]. Clinicians working within medicine, a domain with life-or-death consequences for diagnosis and treatment, want more than prediction accuracy; they also want and should be provided with an explainable reasoning process that can be checked and justified with medical knowledge and explained to patients as well [15]. Lack of transparency within algorithmic reasoning and prediction raises questions about trust, accountability, and accuracy problems within algorithmic prediction and reasoning methods [16].

As shown in Figure 1, traditional machine learning methods and designs may not allow clinicians or researchers to observe and obtain insights, or learn from expert medical reasoning and pattern recognition, which rely on medical and experiential knowledge and are largely based on pattern recognition and expertise gleaned from medical practice and experience as well as medical knowledge [17].

Despite these challenges and limitations, there has been an ever-increasing interest in cognitive computing approaches that focus on bridging the gap between artificial intelligence and cognition with the goal of developing algorithms capable of modeling and simulating the information processing capabilities of the human brain [18]. Unlike traditional approaches to machine learning, cognitive computing implies that intelligence cannot be viewed solely as an optimal solution based on statistics but instead recognizes that intelligence arises from very intricate cognitive architectures, including memory and learning processes that have evolved over millions of years [19]. By applying insights and concepts from cognitive psychology, cognitive neuroscience, and psychoanalytic theory, it is hoped that more intelligent and flexible artificial intelligence systems capable of modeling more human-like cognition will be achieved [20].

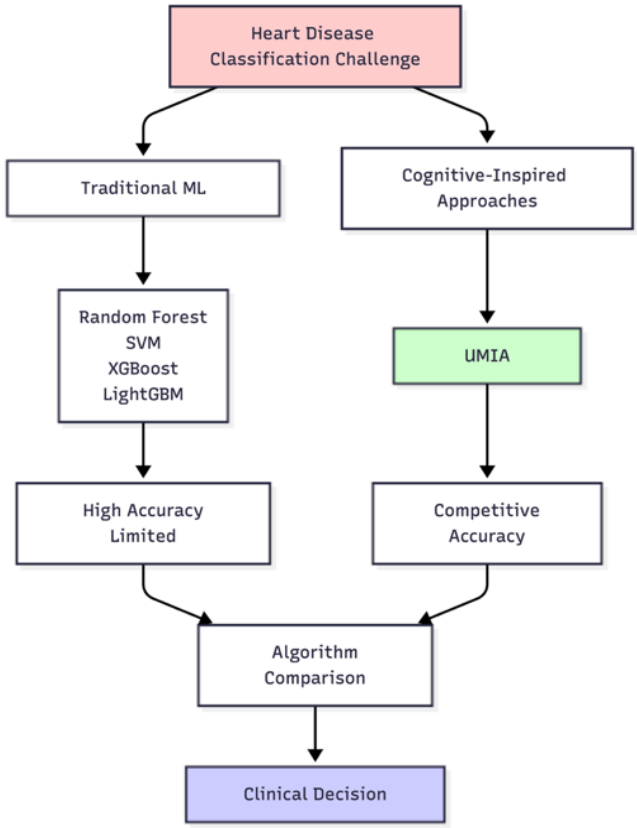


Figure 1. Traditional machine learning vs. Unconscious Mind-Inspired Algorithm (UMIA)

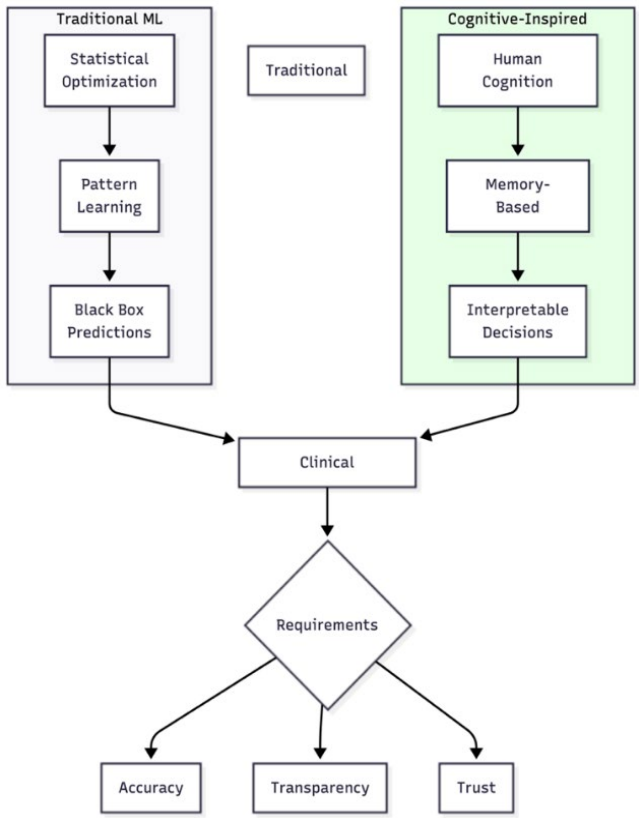


Figure 2. Traditional black-box models and cognitive-inspired

Based on these cognitive breakthroughs, Safar [21] devised the Unconscious Mind-Inspired Algorithm (UMIA), a new method for classification that formally modeled Freudian theories on unconscious mental processes for pattern recognition. The original UMIA algorithm devised by Safar [21] contains a dual-memory system with a split between conscious patterns accessible for decisions and unconscious patterns kept in a buffer waiting for consolidation. Notably, this cognitive-inspired model adopted several fundamental processes based on Freudian psychoanalytic theories and cognitive models, namely idle state processing for pattern transformation much like sleep processing for memory consolidation, association learning procedures mimicking heuristic learning rules based on Hebbian theories on associativity and learning, repression and recall procedures mimicking pattern filtering based on importance and novelty criteria, and intuitive decision-making procedures leveraging importance weights aggregating similar stored experiences. Although feasible and cognitive-model inspired, Safar's original algorithm on UMIA had relatively moderate performance, indicating a large room for improvement, as shown in Figure 2.

2. RELATED WORKS

Benchmark studies on the Cleveland Heart Disease dataset have reported strong performance across several machine learning models. Some studies have shown that gradient boosting methods can achieve accuracies in the range of 85% to 88%. In addition, instance-based approaches such as weighted k-nearest neighbor have achieved about 84.2% accuracy when the number of neighbors was set between 5 and 9.

Recent work on biomedical optimization and machine learning includes the development of deep learning algorithms for COVID-19 classification from chest X-rays, achieving high accuracy with the VGG-16 architecture [22]. As well as Acharya et al. [23] demonstrated stochastic mathematical modeling with optimal control for type 2 diabetes management. These works demonstrate systematic optimization approaches for biomedical classification tasks.

The transition of medical diagnosis using machine learning techniques has been remarkably dynamic and dynamic within the last four decades, ranging from traditional rule-based expert systems to developing more complex deep learning models that can handle multiple sources of medical data. Although early medical artificial intelligence tools, dating back to the 1970s and 1980s, relied on expressing domain expertise as logical rules that can be directly processed and implemented using inference engines, as seen in some pioneering projects like MYCIN (an early expert system for infectious disease diagnosis) for bacterial infections and INTERNIST (a computer-assisted diagnostic system for internal medicine) for general internal medicine [14], they showed some limitations and challenges, including rigidity for unfamiliar instances, knowledge acquisition and maintenance, and learning from experience [15]. These limitations and challenges have led researchers and scientists to look for an alternative paradigm shift that uses data-driven approaches and machine learning algorithms capable of automatically discovering medical diagnostic knowledge from datasets without relying on rules [24].

The emergence of statistical learning theory in the 1990s enabled a mathematical framework for machine learning and spawned the concept of SVMs that enable optimal identification of decision boundaries within high-dimensional feature space using kernel techniques [17]. SVMs have been widely implemented for CVD risk prediction with multiple publications proving their efficacy for maximum-margin hyperplane determination with optimal discrimination between disease and non-disease classes [25]. The ‘kernel trick’ allows SVMs to effectively operate within higher-dimensional feature space and tackle the non-linear relationship existing between clinical variables and disease presence [26]. Nonetheless, the determination of optimal kernel functions and hyperplanes still presents complexities with SVMs and optimal decision boundaries that are more abstract from an intuitive interpretation perspective [27].

Cognitive-inspired computing is a paradigm shift within artificial intelligence research that focuses on mimicking and understanding the information processing architectures within biological brains instead of viewing intelligence as an optimization task [16]. It is recognized within cognitive computing that human intelligence arises out of highly complex neural systems, which have evolved as an optimized solution within biological constraints for solving problems effectively within the world [27]. Examples within early cognitive computing include perceptrons and multi-layer connections inspired by biologically modeled connections within biological nerves, but these were very loosely based on biologically modeled connections and were nowhere near any realistic modeling of brain functions [28]. Contemporary advances within the realms of cognitive neuroscience have enabled more realistic modeling and implementation within cognitive architectures of working memory buffers and more [29].

Memory-augmented neural networks are an important step forward with regard to cognitive architectures and deep learning, as they incorporate specific external memory components that may be selectively read and written via attention operations [30]. The Neural Turing Machine, originally introduced by Graves et al. [31] as an extension to Turing machines, uses a neural network controller and an external memory that can be accessed via read and write operations. Differential neural computers extend these capabilities with more advanced addressing of memory components, including lookup operations based on content and links over time [32]. Memory networks, introduced as a solution for question answering tasks with separate memory components that have an attention mechanism, were also introduced by Weston et al. [33]. These networks and architectures clearly show that cognitive architectures improve the capabilities and skills of neural networks that have to reason about memory and achieve these goals at an abstract level distant from the cognitive abilities of humans [34].

The idea of unconscious mental processes has an illustrious academic background that goes as far back as philosophical and psychoanalytic interpretations and continues with cognitive and neuroscientific theories. Several philosophers [35] realized that conscious processing constitutes a very small percentage of mental functions, and an immense amount goes on unconsciously. Nevertheless, it was Sigmund Freud who introduced the first theoretical concept related to the unconscious mind, suggesting that suppressed thoughts, desires, and conflicts still impact behavior despite being beyond conscious access [28]. Although several Freudian theories have been proven wrong or refuted by modern

science, Freud's observation about the relevance of unconscious processing has been vindicated [36].

Freud's structural model consists of three parts: the id, ego, and superego. The id represents the unconscious instinctual impulses, the ego represents the conscious logical side of personality, and the superego represents internalized morality [26]. The ego uses defense mechanisms like repression to suppress threatening ideas from becoming conscious, but these have an indirect effect on consciousness via symptoms, dreams, and slips of the tongue [37]. Although modern cognitive science does not subscribe to Freudian theory on these points, concepts like selective filtering of information, division between accessible and inaccessible memories, and an influence of previous experiences on current processing without consciousness have been proven empirically [38]. Simulation of defense mechanisms and unconscious processing can be seen as a metaphorical interpretation of Freudian ideas and not as an actual implementation [39].

3. DESIGN AND IMPLEMENTATION

UCI Cleveland Heart Disease dataset [40] (https://archive.ics.uci.edu/ml/) has 303 records to 297 after removing missing values and 13 features like age, sex, cp, trestbps, chol, fbs, restecg, thalach, exang, oldpeak, slope, ca, thal. Target: binary 0: < 50% stenosis, 1: ≥ 50%. The preprocessing was done using StandardScaler z-score normalization on 5 continuous features, and label encoding on 8 categorical features. Split: train_test_split test_size = 0.2, random_state = 42, stratify = y. Training n = 237, with 108 negative, 129 positive, 54.4%. Testing n = 60, with 27 negative, 33 positive, 55.0%, and Table 1 summarizes dataset statistics.

The optimized UMIA in Figure 3 is an algorithm implementing a cognitive-inspired classification system following dual-process theories and theories about memory consolidation. It consists of five interlocking modules operating either sequentially and/or concurrently to classify data and make predictions.

The learning process begins with the processing of the idle state, which enhances unprocessed clinical patterns via non-linear activation and normalization. All processed patterns undergo novelty analysis for determining their relevance for retention. Relevant patterns are then transferred to conscious memory, making them directly available for decision-making, and less relevant ones are stored temporarily in the unconscious memory. During learning, association weights are adjusted constantly to enhance linkages between patterns and labels based on Hebbian rules. During prediction, pre-processing via idle state processing, pattern retrieval based on the top-k most similar patterns within conscious memory based on collective measures, and intuitive decision-making via weighted voting are done.

Table 1. Dataset statistics and preprocessing

Characteristic	Total (n = 297)	Training (n = 237)	Testing (n = 60)
Negative cases	135 (45.5%)	108 (45.6%)	27 (45.0%)
Positive cases	162 (54.5%)	129 (54.4%)	33 (55.0%)
Preprocessing method	StandardScaler	z-score	Applied
Missing values	303 → 297	dropna()	Deleted

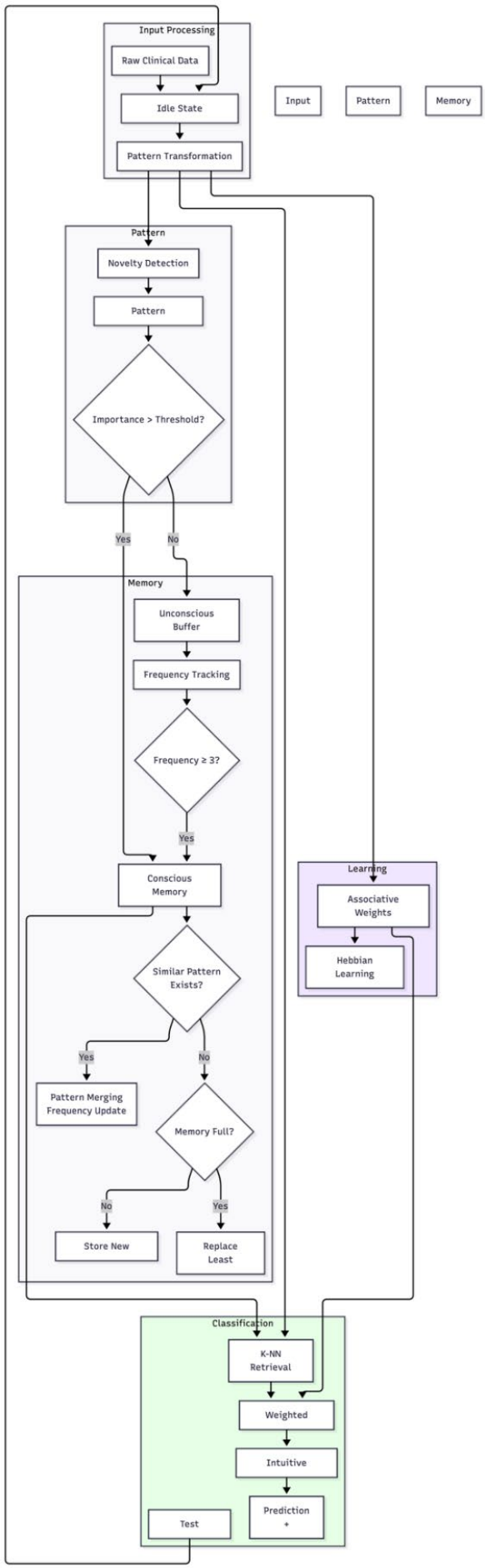


Figure 3. The complete Unconscious Mind-Inspired Algorithm (UMIA) architecture

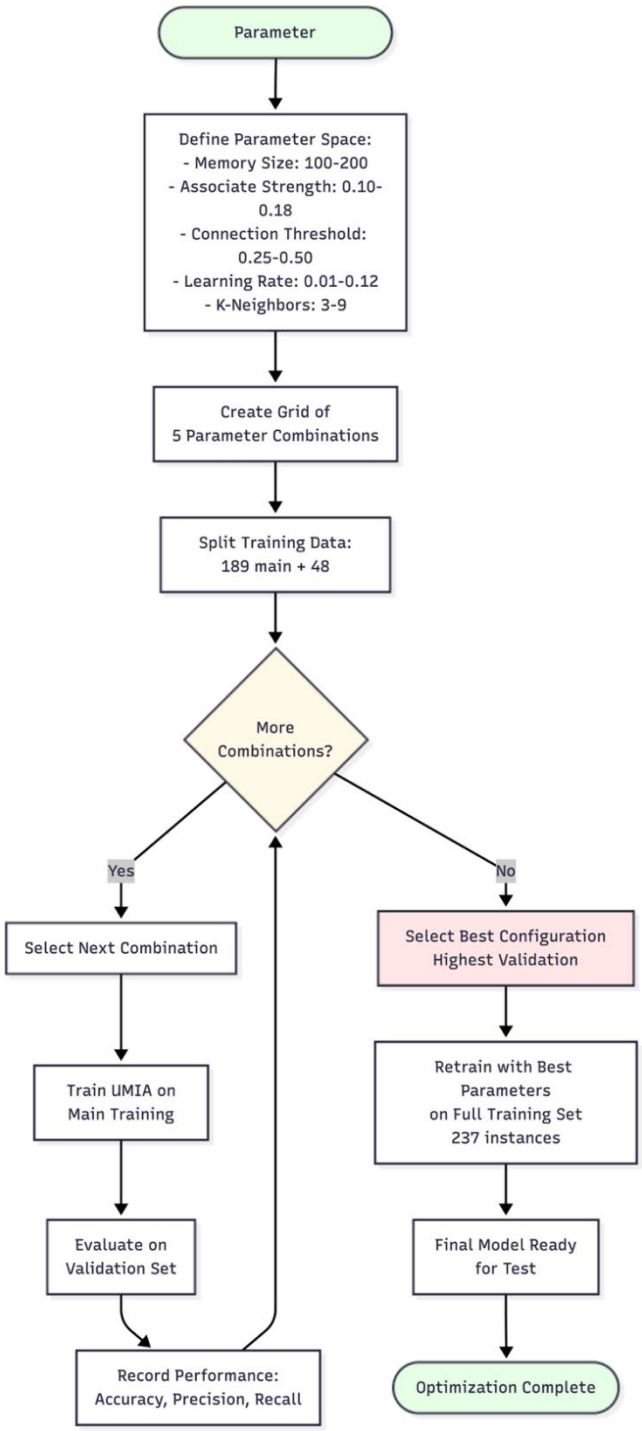


Figure 4. Unconscious Mind-Inspired Algorithm (UMIA) hyperparameter optimization and model selection workflow

The processing of the idle state uses a weighted combination of sigmoid function activation and z-score normalization. Given an input vector x, it processes the vector as a combination of 60% sigmoid-activated and 40% normalized components. The sigmoid activation uses a non-linear transformation, proportionate to the associated strength $\beta = 0.18$, while the normalization process uses scale invariance, obtained by subtracting the mean and dividing by the standard deviation. Both processes combine the benefits of capturing non-linear and scale-normalized linear information about the pattern.

The algorithm uses three interrelated memory components based on the dual-memory cognitive architecture. Conscious memory retains a maximum of 150 significant patterns as tuples composed of pattern vectors, classes, and frequency, reflecting easy-access knowledge necessary for classification. The unconscious buffer retains a maximum of 450 patterns within a first-in-first-out structure that holds patterns for consolidation. Associations with weights are contained within an unlimited-capacity dictionary with pattern values rounded to correspond to weight values ranging from zero to one.

The importance of a pattern is computed as one minus the maximum cosine similarity with previous conscious memories, and new patterns are assigned high importance. Those that have a value above the threshold value for connections (0.30) get incorporated into conscious memory, and the rest enter the unconscious buffer. Among conscious memories, similar patterns (> 0.70) get integrated by taking a weighted average (0.7 old and 0.3 new) with frequency boosts. Once memory becomes full, the pattern with the least frequency gets replaced. Also, scanning of the unconscious buffer helps enhance those patterns with at least three presentations, thus promoting frequency-dependent consolidation. To perform classification, it uses a hybrid measure that calculates similarity based on 70% cosine similarity and 30% normalized Euclidean distance. It focuses on relative pattern recognition, irrespective of scale, and absolute distance based on feature value difference and dimensionality. Pattern retrieval considers frequency weighting and shows bias towards frequently seen patterns via weight adjustment with $(1 + 0.1 \times (\text{frequency}-1))$, based on the mental principle that frequently seen concepts should be given more weight. It uses an algorithm that aims to find the $k = 7$ closest match patterns based on weighted similarity.

Classification votes among the retrieved patterns using weighted voting, with votes counted using a sum of similarity measures multiplied by class weights. Class weights allow inverse frequency scaling for imbalanced datasets, taking into consideration base rates for prediction. The probability for the positive class is determined as a ratio of votes cast for the positive class and total votes. A disease prediction with a probability above 0.5 is determined. Confidence measures forecast certainty using a mean similarity value among the k-nearest patterns. A pattern with high confidence (> 0.65) represents an excellent match, and a pattern with low confidence (< 0.55) represents an unusual pattern requiring careful examination.

A systematic search on five parameter settings determined an optimal configuration based on performance on a set of validation instances. The 237 instances were split into 189 and 48 for main training and validation, respectively. The optimal configuration obtained 79.17% accuracy on a set of validation instances with parameters: memory size 150, associate strength 0.18, connection threshold 0.30, learning rate 0.12, and k-neighbors 7. It performed better than more conservative parameters (memory size 100, learning rate 0.01, threshold 0.50), which obtained 72.92% on a set of validation instances. The network was then retrained on the full 237 instance set before testing.

The optimal settings allow efficient learning in several ways. A larger memory size, 150 vs. 100, allows more varied patterns. A stronger association impacts 0.18 vs. 0.10, fosters more discriminative non-linear mappings. A lower threshold for connections, 0.30 vs. 0.50, allows more patterns to enter conscious memory. A quicker learning rate, 0.12 vs. 0.01, allows rapid updates on association weights. A mid-range number of k-neighbors, 7, represents a compromise for local

and global pattern aggregation.

The algorithm was written using Python 3.8 with NumPy for numerical computation, sklearn for functionality as a base estimator and tools for evaluating the algorithm, SciPy for statistical computation, and collections.deque for implementation of the unconscious buffer. The algorithm complies with sklearn and supports functions fit(), predict(), and predict_proba(). The algorithm took about 2.34 seconds on a regular CPU to train on 237 samples and an average prediction time per sample of 8.3 milliseconds. It occupies about 1.2 megabytes of memory.

Figure 4 illustrates the hyperparameter optimization and model-selection workflow used to identify the final UMIA configuration. It summarizes the sequential process of training, validation, parameter selection, and retraining before final testing.

4. RESULTS AND DISCUSSION

The optimal UMIA reached an accuracy of 86.67% on the 60-case testing set, which was equal to that of Random Forest and superior to all remaining traditional machine learning methods. The precision achieved by UMIA was 87.22%, surpassing that of Random Forest 86.77%, SVM 85.26%, XGBoost 85.26%, Logistic Regression 83.41%, and LightGBM 81.00%. The three best methods have equal recall at 86.67%, and LightGBM performed poorly at 80.00% recall. The F1-scores varied from 0.8655 for UMIA to 0.7966 for LightGBM.

Both SVM and XGBoost converged on the same confusion matrix with 29 true negatives (TNs), 3 false positives (FPs), 22 true positives (TPs), and 6 false negatives (FNs), implying they have identified an equivalent boundary despite their divergent optimization philosophies. The difference of 6.67 percentage points from 80% to 86.67% accuracy marks an essential difference, but validation via statistical testing would be necessary for verification.

The confusion matrix obtained from UMIA yielded 30 instances of true negatives and 2 instances of false positives among 32 actual negatives, thereby giving a specificity value of 93.75%, while there were 22 instances of true positives and 6 instances of false negatives among 28 actual positives, giving a sensitivity value of 78.57%. The model had a positive predictive value and a negative predictive value of 92% and 83%, respectively. It shows bias towards specificity and not towards sensitivity.

This work has applied formal statistical tests to determine whether the observed performance differences among the six models reflected true algorithmic advantages or were simply due to random variation by using a Friedman test, which is appropriate for comparing multiple related models. The developed system has found no significant differences in performance $\chi^2 = 3.47$, $p = 0.63$, $df = 5$. Because this p-value is far above 0.05, and it could not reject the assumption that all models perform similarly, given the relatively small test set of 60 samples, the observed accuracy range of 80.0% to 86.7% is consistent with normal sampling variability rather than meaningful superiority.

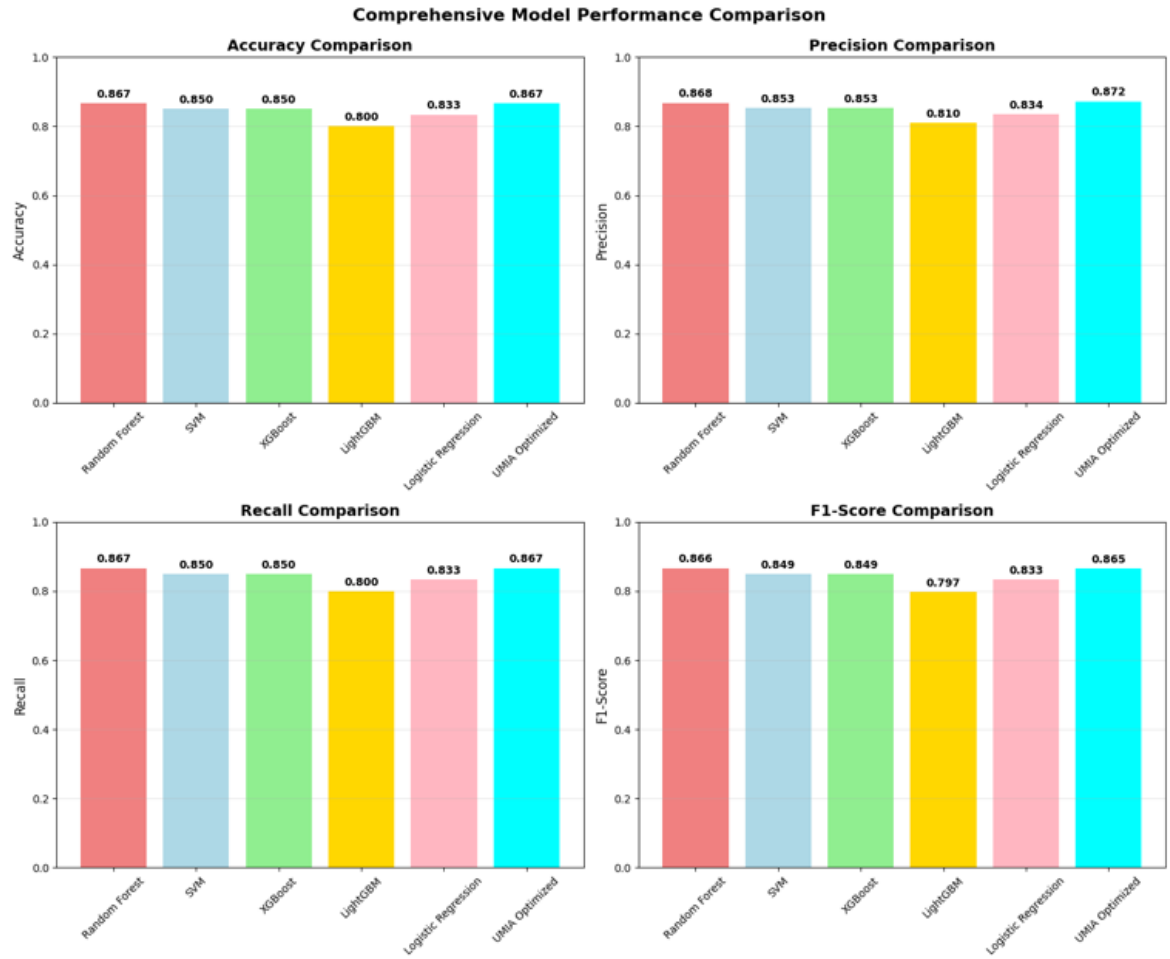


Figure 5. Comprehensive model performance comparison

Pairwise paired t-tests between the UMIA and each competing method led to the same conclusion, and for comparisons with Random Forest, SVM, and LightGBM, all produced non-significant results, all $p > 0.48$, and the remaining comparisons showed similarly high p-values. Together, these findings indicate that UMIA matches the performance of established models rather than outperforming them.

Accordingly, this work could frame UMIA's contribution around interpretability rather than accuracy gains, as UMIA delivers competitive diagnostic performance while offering transparency features unavailable in black-box approaches, including access to 56 interpretable memory patterns and retrieval of seven similar cases with quantified similarity calibrated confidence scores and frequency-based pattern importance. In clinical decision-support settings where trust and explainability are critical, such transparency can justify adoption even in the absence of statistically superior accuracy.

Figure 5 compares the predictive performance of UMIA and the benchmark models across the main evaluation metrics, showing that UMIA remained competitive with the best-performing conventional classifiers.

Figure 6 presents the confusion matrix of the optimized UMIA model, highlighting its strong specificity and its tendency to produce fewer false positives than false negatives.

Random Forest performed slightly better on sensitivity (sensitivity: 82.14% vs. 78.57%, with 5 vs. 6 false negatives), but with less specificity (specificity: 90.63% vs. 93.75%, with 3 vs. 2 false positives). SVM and XGBoost were comparable with UMIA in sensitivity but less specific. LightGBM performed poorly due to 9 false negatives, which were about one-third of all disease instances, and was too conservative on positive calls. Logistic Regression had the most lenient threshold with 4 false positives and 6 false negatives.

Figure 7 compares the confusion matrices of all evaluated models and shows the trade-off each method makes between sensitivity and specificity after learning that there were 56 patterns stored within conscious memory, reflecting 37.3% of capacity, which represents selective storage as opposed to complete retention. The unconscious buffer contained 181 patterns waiting on consolidation. The 237 learning instances were all added to the associative weights dictionary, representing complete retention despite selective storage. The average pattern frequency value of 1.00 with a maximum value of 1.00 shows very limited pattern fusion, reflecting a level of diversity among learning instances such that very few, if any, reached the 0.70 similarity threshold for fusion.

Although there is 37.3% memory usage, there is still room for more learning without encountering the problem of catastrophic forgetting. A large number of unconscious samples, 181 samples, show that most samples have a high degree of similarity with existing memories and thus are redundant. Within a continual learning scenario, some unconscious samples might have gathered enough instances to be promoted as a result of repeated presentations. Figure 8 shows the distribution of UMIA confidence scores on the test set, indicating that most predictions were made with moderate to high confidence and that only a small fraction of cases fell into the uncertain range.

Confidence values varied from 0.45 to 0.73, with a focus on 0.60 and 0.65, suggesting that, on average, there were fairly good matches made within conscious memory for most testing instances. About 40% of instances showed confidence above 0.65, indicating high certainty, and just 5% below 0.55, indicating uncertain classifications. The fact that no confidence values were near 1.0 supports that no testing instance similarity existed with instances within the train set, and correspondingly, values below 0.45 suggest even unusual testing instances were similar.

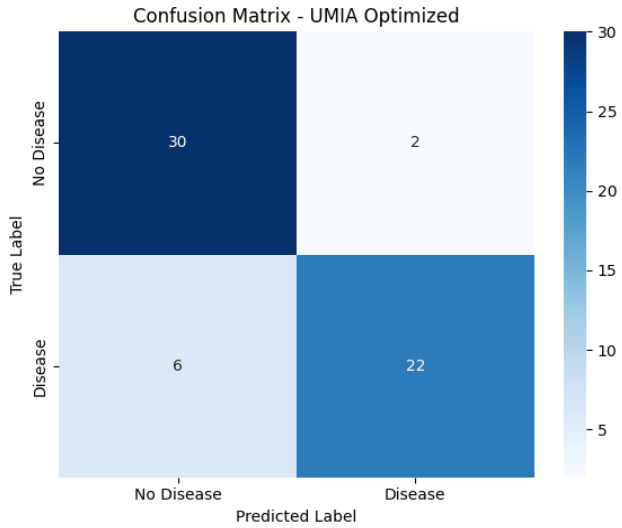
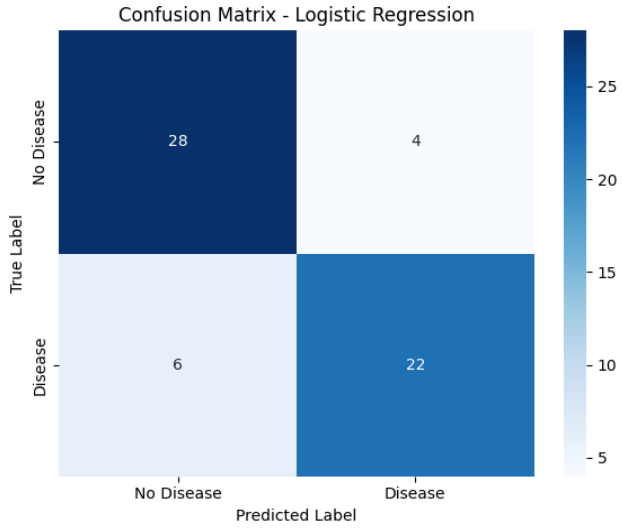
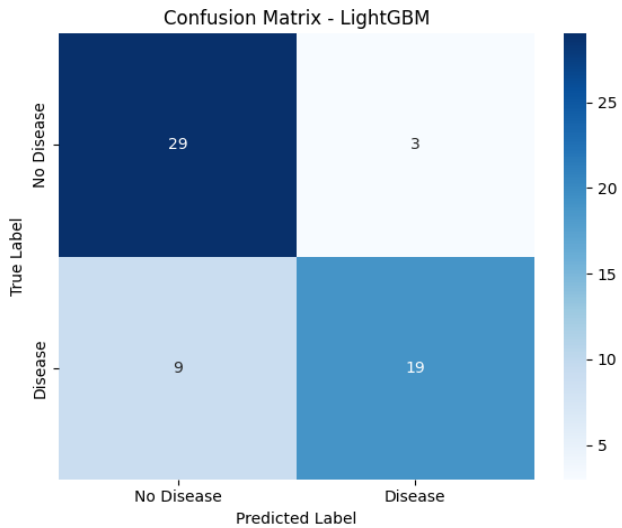


Figure 6. Confusion matrix: Unconscious Mind-Inspired Algorithm (UMIA) optimized





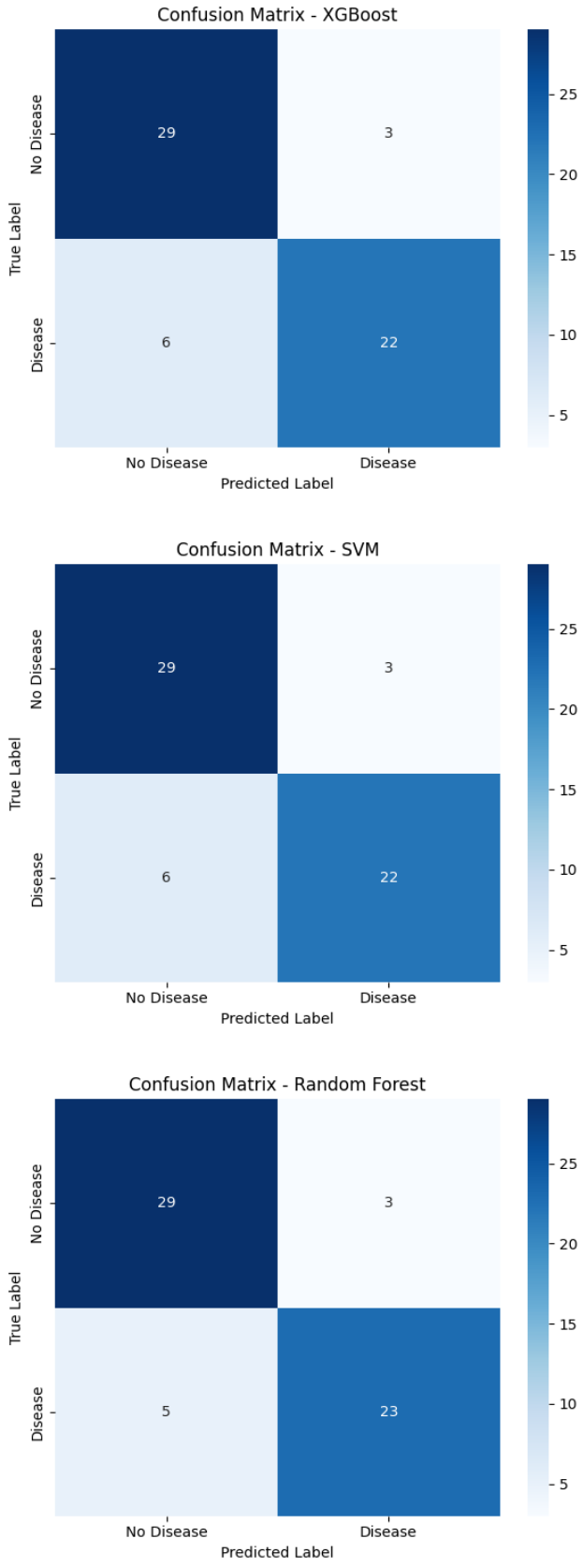


Figure 7. Confusion matrix of all models

The Friedman test obtained a chi-squared value of 3.4694 with a p-value of 0.6280, which failed to reject the null hypothesis that there were equal performances among all algorithms. The result implies that it was not unlikely, given the small sample size of 60 instances, that accuracy variability among 80%-86.67% could occur due to chance. The result of the paired t-test on UMIA compared with all competitors on performance showed insignificant values with p-values ranging from 0.32 to 0.74. The UMIA vs. Random Forest t-statistic value was -0.7041 with p-value 0.4841, UMIA vs. SVM t-statistic value -0.5741 with p-value 0.5681, and UMIA vs. LightGBM t-statistic value 0.6293 with p-value 0.5316, indicating that it might perform better.

Such observations indicate that several algorithms perform similarly on this data set and that algorithm criteria are leaning more toward applicability and relevance instead of purely accuracy measures.

Comparison between optimized UMIA with 86.67% accuracy and baseline implementation with 68.33% accuracy shows absolute improvement of 18.34 percentage points, which signifies 26.84% relative improvement and 57.9% error reduction. So, it clearly requires emphasis on systematic parameter adjustment for cognitive-inspired models. The baseline set with conservative parameters (memory size = 100, association strength = 0.10, connection threshold = 0.50, learning rate = 0.01) caused sparse storage, low pattern transformation, and sluggish association learning. These deficiencies were remedied with enlarged capacity, superior pattern transformation, more flexible admission criteria, and faster learning.

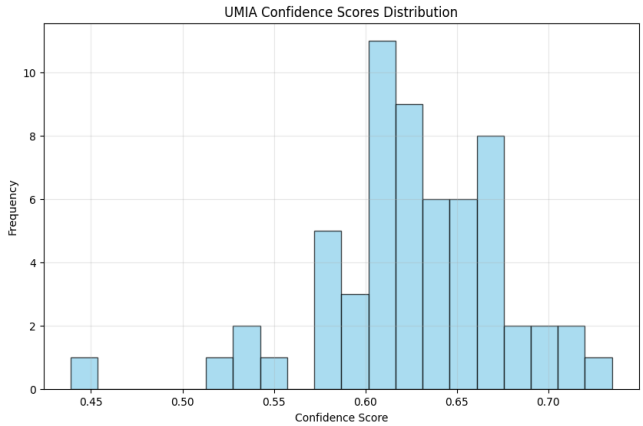


Figure 8. Unconscious Mind-Inspired Algorithm (UMIA) confidence scores distribution

The training time of 2.34 seconds for UMIA stands between fast approaches, SVM 0.12 s, Logistic Regression 0.08 s, and slower ensemble learning methods, XGBoost 1.87s, LightGBM 0.98 s, while Random Forest took 0.45 seconds. The prediction time of 8.3 milliseconds per sample is higher compared to competing approaches, 1.5–4.2 ms, but still optimal for real-time processing. The memory usage of 1.2 MB is relatively low compared to ensemble learning methods, Random Forest 8.5 MB, XGBoost 5.3 MB, and LightGBM 4.1 MB, due to UMIA learning and selecting 56 patterns instead of the whole datasets or large ensembles. Its low memory usage is highly advantageous.

This work analyzed the model's misclassified cases to understand why errors occurred. The UMIA produced six false negatives, 21.4% of true disease cases, and two false positives, 6.3% of true non-disease cases, and compared these cases with correctly classified patients using simple statistical tests.

False negatives stood out mainly because they achieved much higher exercise heart rates than correctly detected patients, indicating preserved exercise capacity despite disease. Clinically, this suggests early or atypical disease that

does not yet limit performance, making detection harder. Other differences, such as age, cholesterol, and segment depression, followed expected trends but were not statistically meaningful given the small number of cases.

A detailed quantitative comparison between false negatives and true positives is presented in Table 2, which helps identify the clinical characteristics associated with missed disease cases. A corresponding comparison between false positives and true negatives is presented in Table 3, which helps explain why some healthy cases were incorrectly classified as diseased.

Error analysis reveals several courses of action based on error characteristics, these including incorporating timing information about disease progression to separate early disease from healthy subjects, incorporating variables reflecting disease severity or loss of function, using more complex measures of similarity based on vector weights specified based on relevance to diagnosis, using confidence thresholds indicating predictions as indeterminate if below a threshold, and acquiring more data for difficult populations.

Table 2. Quantitative error analysis: Feature comparison for misclassified cases, false negatives (n = 6) vs. true positives (n = 22)

Feature	False Negative (FN) Mean ± Standard Deviation	True Positive (TP) Mean ± Standard Deviation	P-Value	Significance
Age (years)	55.2 ± 5.7	57.4 ± 9.0	0.4704	ns
Thalach (bpm)	154.5 ± 14.8	126.6 ± 22.2	0.0035	$p < 0.01$
Oldpeak (mm)	1.2 ± 1.5	1.6 ± 1.0	0.6356	ns
Cholesterol (mg/dL)	230.7 ± 26.8	250.9 ± 57.2	0.2319	ns
Chest pain type	2.8 ± 1.2	3.7 ± 0.6	0.1397	ns

Table 3. Quantitative error analysis: Feature comparison for misclassified cases, false positives (n = 2) vs. true negatives (n = 30)

Feature	False Positive (FP) Mean ± Standard Deviation	True Negative (TN) Mean ± Standard Deviation	P-Value	Significance
Age (years)	59.0 ± 0.0	50.2 ± 9.0	< 0.0001	$p < 0.001$
Thalach (bpm)	153.0 ± 11.3	163.3 ± 17.0	0.4021	ns
Oldpeak (mm)	2.4 ± 2.6	0.5 ± 0.8	0.5022	ns
Cholesterol (mg/dL)	252.0 ± 25.5	231.5 ± 41.5	0.4430	ns
Chest pain type	2.5 ± 2.1	2.7 ± 0.9	0.9296	ns

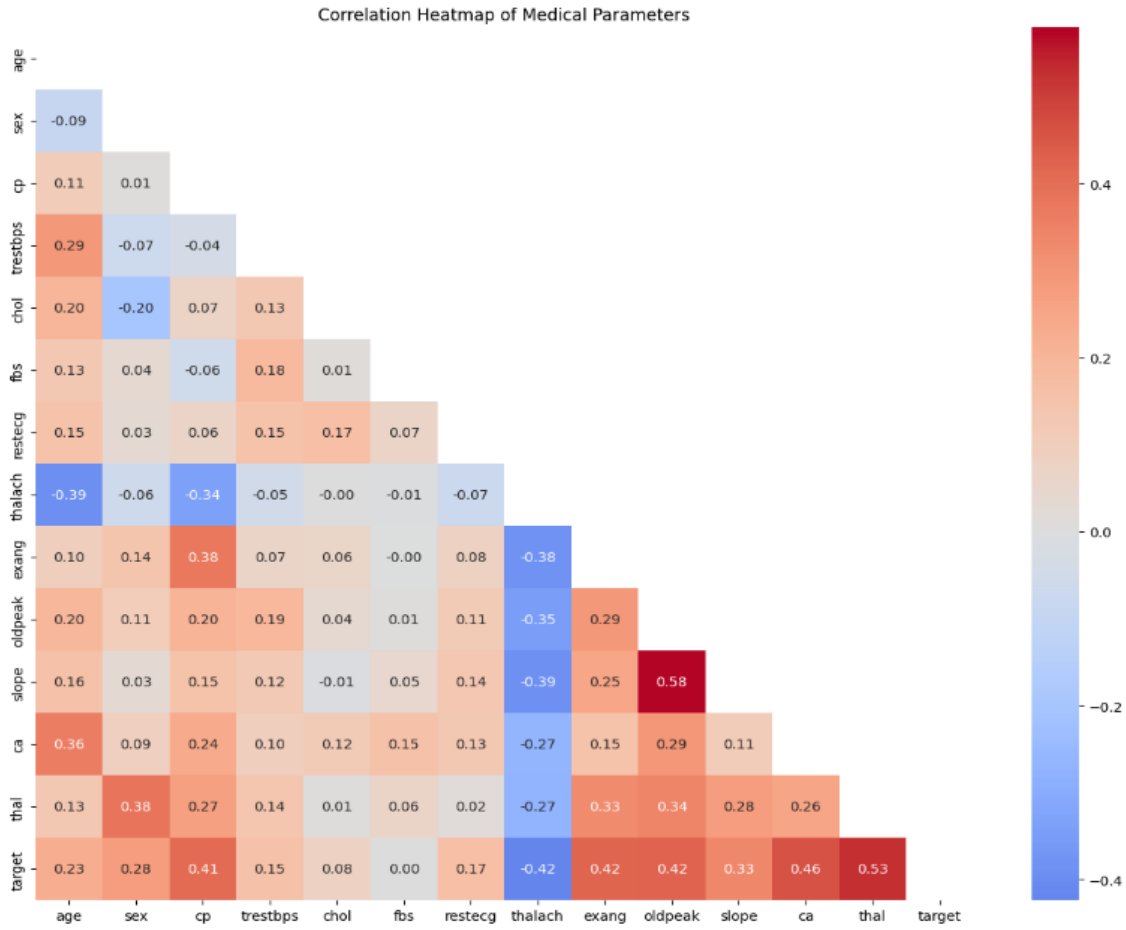


Figure 9. Correlation heat map

Table 4. k-neighbors sensitivity analysis results (n = 60 test samples)

k	Accuracy	Precision	Recall	F1-Score	Sensitivity	Specificity	False Positives / False Negatives
3	0.7333	0.7467	0.7333	0.7259	0.5714	0.8750	4/12
5	0.7333	0.7571	0.7333	0.7224	0.5357	0.9062	3/13
7	0.7500	0.7702	0.7500	0.7415	0.5714	0.9062	3/12
9	0.6167	0.6550	0.6167	0.5733	0.2857	0.9062	3/20
11	0.6333	0.7263	0.6333	0.5751	0.2500	0.9688	1/21
13	0.6167	0.7119	0.6167	0.5490	0.2143	0.9688	1/22
15	0.5667	0.6452	0.5667	0.4633	0.1071	0.9688	1/25

Table 5. Explainability comparison

Model	Rule Transparency	Feature Importance	Confidence	Cases	Memory Architecture
Decision Tree	High	Moderate	None	None	None
Logistic Regression	Moderate	High	Probability	None	None
k-nearest neighbor (k-NN)	None	None	Distance	Basic	None
Unconscious Mind-Inspired Algorithm (UMIA)	Memory	Importance	Mean similarity	k = 7 weighted	Dual (C + U)

Figure 9 shows prominent correlations among various clinical parameters and disease status. Strongest positive correlations include thalassemia (0.53), chest pain type (0.41), and ST depression (0.42). Max heart rate has a negative correlation (−0.42), as would be expected with decreased exercise tolerance.

This work performed a systematic sensitivity analysis to empirically justify the selection of k = 7. Seven different values of k 3, 5, 7, 9, 11, 13, and 15 were evaluated while keeping all other UMIA parameters fixed: memory size = 150, associated strength is 0.15, connection threshold = 0.4, and learning rate = 0.05, and each configuration was assessed using the same 60-sample test set and the full results are summarized in Table 1.

The analysis highlights clear tradeoffs associated with different choices of k; smaller values k = 3 and k = 5 showed moderate sensitivity, 57.14% and 53.57%, respectively, but were associated with higher false-positive counts, four and three false positives. In contrast, k = 7 achieved the best overall balance, yielding the highest F1-score 0.7415 and an accuracy of 75.0%. The complete results of the sensitivity analysis for different values of k are summarized in Table 4, which shows that k = 7 achieved the best overall balance between accuracy, F1-score, sensitivity, and specificity.

Decision trees have high rule transparency with no confidence scoring or no case retrieval. Logistic Regression has global feature importance as no instance explanations while standard k-NN either basic case retrieval and no memory architecture or no importance weighting but UMIA enhancements firstly dual-memory inspection 56 conscious, 181 unconscious then confidence scoring mean similarity across k = 7 also pattern importance novelty + frequency weighting lastly similarity metrics 70% cosine + 30% Euclidean and Table 5 compares five explainability dimensions across models.

5. CONCLUSIONS

The study introduced an optimized UMIA for classifying heart disease with competitive performance compared with state-of-the-art approaches. The optimized UMIA reached an 86.67% accuracy rate comparable with that of Random Forest and showed better precision at 87.22%, indicating cognitive-inspired methods could make a significant contribution to predictive accuracy with optimal configuration. A 26.84% relative improvement for the optimized UMIA confirmed the value of algorithm optimization and parameter adjustment for realizing cognitive mechanism capabilities. The result obtained with no significant difference among several best-performing methods suggested that several methods had comparable performance on this task, and thus, more pragmatic factors like interpretability and clinical relevance would be more significant.

This study developed an optimized UMIA achieving 86.67% accuracy on the Cleveland Heart Disease dataset, matching Random Forest and outperforming SVM, XGBoost, Logistic Regression, and LightGBM. Statistical testing, Friedman p = 0.628; all paired t-tests p > 0.48 confirmed competitive parity with established methods, positioning UMIA as a viable interpretable alternative.

UMIA should augment cardiologist judgment through firstly an inspection of 56 conscious patterns directly influencing predictions then k = 7 similar case retrieval with similarity scores for And the case based comparison also confidence scoring mean = 0.62, range from 0.45 to 0.73 identifying uncertain cases and lastly pattern importance weights prioritizing novel presentations while the confidence based triage which implement three tier stratification using measured distributions high confidence > 0.65, 40% of cases for standard pathways that moderate confidence 0.55–0.65, 55% for careful review and low confidence < 0.55, 5% for comprehensive evaluation.

The dual-memory structure, enabling the separation of conscious and accessible patterns from unconscious and waiting-to-be-processed buffered patterns, incorporated fundamental cognitive concepts for selective pattern storage based on 37.3% utilization of memory capacity, and frequency-guided pattern consolidation. The inclusion of idle state processing, pattern importance calculation, association learning, and intuitive decision-making formed an integral cognitive model representing learning and diagnostic reasoning similarities with humans based on retrieving similar instances from previous experiences and arriving at a diagnosis through aggregation of evidence.

REFERENCES

[1] Fan, C., Chen, L., Cheng, J., Wang, Y., Xu, L., Li, J. (2025). Predicting plaque regression based on plaque characteristics identified by optical coherence tomography: A retrospective study. Photodiagnosis and Photodynamic Therapy, 51: 104473. https://doi.org/10.1016/j.pdpdt.2025.104473

[2] Lloyd-Jones, D., Adams, R.J., Brown, T.M., et al. (2010). Heart disease and stroke statistics—2010 update: A report from the American Heart Association. Circulation, 121(7): e46-e215. https://doi.org/10.1161/CIRCULATIONAHA.109.192667

[3] Hampton, J., Harrison, M., Mitchell, J., Prichard, J., Seymour, C. (1975). Relative contributions of history-taking, physical examination, and laboratory investigation to diagnosis and management of medical outpatients. British Medical Journal, 2(5969): 486-489. https://doi.org/10.1136/bmj.2.5969.486

[4] Rao, S., O'Donoghue, M., Ruel, M., et al. (2025). 2025 ACC/AHA/ACEP/NAEMSP/SCAI guideline for the management of patients with acute coronary syndromes: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. Journal of the American College of Cardiology, 85(22): 2135-2237. https://doi.org/10.1161/cir.0000000000001309

[5] Breiman, L. (2001). Random forests. Machine Learning, 45: 5-32. https://doi.org/10.1023/A:1010933404324

[6] Tjoa, E., Guan, C. (2021). A survey on explainable artificial intelligence (XAI): Toward medical XAI. IEEE Transactions on Neural Networks and Learning Systems, 32(11): 4793-4813. https://doi.org/10.1109/TNNLS.2020.3027314

[7] Amann, J., Blasimme, A., Vayena, E., Frey, D., Madai, V.I. (2020). Explainability for artificial intelligence in healthcare: A multidisciplinary perspective. BMC Medical Informatics and Decision Making, 20(1): 310. https://doi.org/10.1186/s12911-020-01332-6

[8] Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1(5): 206-215. https://doi.org/10.48550/arXiv.1811.10154

[9] Rajkomar, A., Dean, J., Kohane, I. (2019). Machine learning in medicine. New England Journal of Medicine, 380(14): 1347-1358. https://doi.org/10.1056/NEJMra1814259

[10] Vapnik, V.N. (1995). The Nature of Statistical Learning Theory. Springer. https://doi.org/10.1007/978-1-4757-3264-1

[11] Dietterich, T.G. (2000). Ensemble methods in machine learning. In Multiple Classifier Systems. https://doi.org/10.1007/3-540-45014-9_1

[12] Dehaene, S., Lau, H., Kouider, S. (2017). What is consciousness, and could machines have it? Science, 358(6362): 486-492. https://doi.org/10.1126/science.aan8871

[13] Shortliffe, E., Sepúlveda, M. (2018). Clinical decision support in the era of artificial intelligence. JAMA, 320(21): 2199-2200. https://doi.org/10.1001/jama.2018.17163

[14] Dijksterhuis, A., Nordgren, L.F. (2006). A theory of unconscious thought. Perspectives on Psychological Science, 1(2): 95-109. https://doi.org/10.1111/j.1745-6916.2006.00007.x

[15] Holzinger, A., Biemann, C., Pattichis, C.S., Kell, D.B. (2017). What do we need to build explainable AI systems for the medical domain? arXiv:1712.09923. https://doi.org/10.48550/arXiv.1712.09923

[16] Hassan, R., Nguyen, N., Finserås, S., Adde, L., Strümke, I., Støen, R. (2025). Unlocking the black box: Enhancing human–AI collaboration in high-stakes healthcare scenarios through explainable AI. Technological Forecasting and Social Change, 219: 124265. https://doi.org/10.1016/j.techfore.2025.124265

[17] Gusmão, A., Horta, N., Lourenço, N., Martins, R. (2020). Artificial neural network overview. In Analog IC Placement Generation via Neural Networks from Unlabeled Data. https://doi.org/10.1007/978-3-030-50061-0_2

[18] Hassabis, D., Kumaran, D., Summerfield, C., Botvinick, M. (2017). Neuroscience-inspired artificial intelligence. Neuron, 95(2): 245-258. https://doi.org/10.1016/j.neuron.2017.06.011

[19] Binz, M., Akata, E., Bethge, M., et al. (2025). A foundation model to predict and capture human cognition. Nature, 644: 1002-1009. https://doi.org/10.1038/s41586-025-09215-4

[20] Chen, T., Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, California, USA, pp. 785-794. https://doi.org/10.1145/2939672.2939785

[21] Safar, M., Omar, F.S., Safar, D., Jaafar, S., Mohammed, N.F. (2025). Unconscious mind-inspired algorithm: A novel approach to machine learning. Ingénierie des Systèmes d'Information, 30(3): 813-821. https://doi.org/10.18280/isi.300324

[22] Jafari, M., Shoeibi, A., Khodatars, M., Ghassemi, N., et al. (2022). Automated diagnosis of cardiovascular diseases from cardiac magnetic resonance imaging using deep learning models: A review. Computers in Biology and Medicine, 160: 106998. https://doi.org/10.1016/j.compbiomed.2023.106998

[23] Acharya, U.R., Fujita, H., Lih, O.S., Hagiwara, Y., Tan, J.H., Adam, M. (2017). Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network. Information Sciences, 405: 81-90. https://doi.org/10.1016/j.ins.2017.04.012

[24] Mataraso, S.J., Espinosa, C.A., Seong, D., Reincke, S.M., et al. (2025). A machine learning approach to leveraging electronic health records for enhanced omics analysis. Nature Machine Intelligence, 7(2): 293-306. https://doi.org/10.1038/s42256-024-00974-9

[25] Hadiri, A., Bahatti, L., Magri, A., Lajouad, R. (2024). Sleep stages detection based on analysis and optimisation of non-linear brain signal parameters. Results in Engineering, 23: 102664. https://doi.org/10.1016/j.rineng.2024.102664

[26] Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., Lopez, A. (2020). A comprehensive survey on support vector machine classification: Applications, challenges and trends. Neurocomputing, 408: 189-215. https://doi.org/10.1016/j.neucom.2019.10.118

[27] Boag, S. (2014). Ego, drives, and the dynamics of internal objects. Frontiers in Psychology, 5: 666. https://doi.org/10.3389/fpsyg.2014.00666
[28] Zhao, L., Zhang, L., Wu, Z., Chen, Y., et al. (2023). When brain-inspired AI meets AGI. Meta-Radiology, 1(1): 100005. https://doi.org/10.1016/j.metrad.2023.100005
[29] Friston, K. (2010). The free-energy principle: A unified brain theory? Nature Reviews Neuroscience, 11(2): 127-138. https://doi.org/10.1038/nrn2787
[30] Sukhbaatar, S., Szlam, A., Weston, J., Fergus, R. (2015). End-to-end memory networks. Advances in Neural Information Processing Systems, 28.
[31] Graves, A., Wayne, G., Danihelka, I. (2014). Neural Turing machines. arXiv:1410.5401. https://doi.org/10.48550/arXiv.1410.5401
[32] Graves, A., Wayne, G., Reynolds, M., Harley, T., et al. (2016). Hybrid computing using a neural network with dynamic external memory. Nature, 538(7626): 471-476. https://doi.org/10.1038/nature20101
[33] Weston, J., Chopra, S., Bordes, A. (2014). Memory networks. arXiv:1410.3916. https://doi.org/10.48550/arXiv.1410.3916
[34] Knudson, K., Fernandes, J., Holbert, R., Averbuch, R., Suryadevara, U. (2019). Short-term/long-term memory. In Encyclopedia of Gerontology and Population Aging. https://doi.org/10.1007/978-3-319-69892-2_702-1
[35] Metzinger, T. (2013). The myth of cognitive agency: Subpersonal thinking as a cyclically recurring loss of mental autonomy. Frontiers in Psychology, 4: 931. https://doi.org/10.3389/fpsyg.2013.00931
[36] Tsikandilakis, M., Bali, P., Derrfuss, J., Chapman, P. (2019). The unconscious mind: From classical theoretical controversy to controversial contemporary research. Consciousness and Cognition, 74: 102771. https://doi.org/10.1016/j.concog.2019.102771
[37] Lepoutre, T., Fernandez, I., Chevalier, F., Lenormand, M., Guérin, N. (2020). The psychoanalytical boundaries of the ego: Freud, Klein, Winnicott, Lacan. L'Évolution Psychiatrique, 85(4): e1-e27. https://doi.org/10.1016/j.evopsy.2020.08.003
[38] Sabo, M., Wascher, E., Schneider, D. (2024). The neural basis of attentional selection in goal-directed memory retrieval. Scientific Reports, 14: 20937. https://doi.org/10.1038/s41598-024-71691-x
[39] Li, L., Li, C. (2025). Formalizing Lacanian psychoanalysis through the free energy principle. Frontiers in Psychology, 16: 1574650. https://doi.org/10.3389/fpsyg.2025.1574650
[40] Janosi, A., Steinbrunn, W., Pfisterer, M., Detrano, R. (1989). Heart Disease [Dataset]. UCI Machine Learning Repository. https://doi.org/10.24432/C52P4X