RESEARCH PAPER

Quantum Behaved Particle Swarm Optimization Based Feature Extraction with Adaptive Extreme Learning Machine Classifier for Chronic Kidney Disease and Cardiovascular Disease Prediction

M. Chandra Prabha[1*], Dr. D. Anitha[2]

[1*]Ph.D. Research Scholar Department of Computer Science Sri Ramakrishna College of Arts and Science for Women Coimbatore. prabhaapsara13@gmail.com

[2]Associate Professor Department of Computer Science Sri Ramakrishna College of Arts and Science for Women Coimbatore. anithacs@srcw.ac.in

Abstract:
Early and accurate prediction of Cardiovascular Disease (CVD) and Chronic Kidney Disease (CKD) is important for reducing morbidity through timely clinical intervention. This study proposes a novel Quantum-Behaved Particle Swarm Optimization–Based Feature Extraction with Adaptive Extreme Learning Machine (QPSO–AELM) classifier designed to improve diagnostic performance in multi-disease prediction environments. QPSO is employed to identify the most discriminative feature subsets from high-dimensional clinical data, while the adaptive ELM component confirms robust nonlinear classification and fast convergence. The proposed model was estimated using publicly available CVD and CKD datasets and benchmarked against contemporary approaches, which include SSL-GLP, MSM-CKD, LGBM-CKD, and SHAP-CVD. Experimental validation demonstrated reliably superior performance through enhanced accuracy, specificity, sensitivity, and MCC values, confirming the reliability and strong generalization of the QPSO–AELM framework. The results highlight the capability of combining quantum-behaved optimization with adaptive neural learning to deliver scalable and efficient medical decision-support systems.

Keywords: Quantum-Behaved Particle Swarm Optimization, Adaptive Extreme Learning Machine, Feature Extraction, Cardiovascular Disease Prediction, Chronic Kidney Disease Prediction, Medical Decision Support Systems

How to cite this article: Chandra Prabha M, Anitha D. Quantum Behaved Particle Swarm Optimization Based Feature Extraction with Adaptive Extreme Learning Machine Classifier for Chronic Kidney Disease and Cardiovascular Disease Prediction. Int J Drug Deliv Technol. 2026;16(35s): 673-692. DOI: 10.25258/ijddt.16.35s.76

1. Introduction

Both Cardiovascular Disease (CVD) and Chronic Kidney Disease (CKD) rank as the primary causes of prolonged disability and death, posing animportant burden on healthcare systems globally. According to clinical statistics, CVD continues to account for the highest number of deaths across the world, while CKD often progresses silently, with many patients remaining asymptomatic until advanced stages.Current studies highlight that early diagnosis of CVD and CKD is important for inhibiting organ failure, improving patient survival, and reducing the costs of healthcare. However, the heterogeneity, complexity, and high dimensionality of medical datasets make early prediction extremely challenging, particularly when multiple comorbid factors are involved.

Over the past ten years, research in CVD prediction has witnessed both limitations and advancements. Medication-compliance analytics [12], GA–NN hybrid classification [13], and discriminant-analysis–based risk prediction tools [14] offered early solutions but faced some nonlinear feature interactions. Remote health monitoring systems like Wanda-CVD [15] have proved the value of contextual features but lacked advanced capabilities of feature extraction. High-risk patient focus in dyslipidemia research [16] and Deep attention models for cardiac risk [17] provided useful directions but were computationally rigorous. Time-aware fusion for CVD risk prediction [18], simplified Naïve Bayes–based prediction frameworks [19], and cardiology-focused deep learning review models [20] requireda holistic predictive capability. Even though tree-based models [21] and findings from metabolic-risk studies [22] improved diagnosis, they required long training times, large datasets, and established limited interpretability. Current works on heart disease prediction performance [23], the DASMcC framework [24], integrated machine-learning-based cardiovascular risk prediction frameworks [25], and EHR-driven ML/DL frameworks [26] enhanced prediction accuracy but did not include lightweight adaptive classifiers or evolutionary feature optimization.

Similarly, many computational intelligence techniques have been applied for CKD prediction. Attribute-selection–driven models such as wrapper-subset evaluators [1] and reliability-based progression modeling [2] presented good behaviour, but lacked robust nonlinear feature learning. Evolutionary neural-network approaches like Modified Cuckoo Search–trained NN (MCS-NN) [3] enhanced classification but suffered from sensitivity to initial weights and slower convergence. Linear feature analysis using CSP and LDA [4] successfully identified significant biomarkers but was limited in capturing nonlinear correlations. Optimization-driven approaches, such as ZMPC-based anemia management [5] and XGBoost-based diagnosis

*Author for Correspondence: prabhaapsara13@gmail.com

[6], achieved high accuracy but relied heavily on handcrafted or manually derived features. In the same way, clinical decision support for cardiac risk [7], time-aware deep learning models for CVD risk prediction [8], and tree-based approaches [9] and CVD risk modeling using genetic data [10] contributed useful insights but did not combine feature optimization and classification into a unified framework.Though NIRS-based outcome prediction verified clinically valuable, its interpretation still depended on expert clinical judgment [11].

2. Related Works

The story of predicting CVD and CKD has unfolded over a decade of restless experimentation, each study trying to read the body's hidden warnings a slight earlier and clearer. Researchers began by wrestling first with attributes themselves, attempting to understand which clinical clues truly mattered. Chetty et al. showed that thoughtful attribute selection sharpens CKD classifiers and cuts through diagnostic noise [1], while Nishanth and Thiruvaran affirmed that serum creatinine, specific gravity, albumin, and hemoglobin hold a quiet but significant voice in early detection of CKD [4]. This growing attention to discriminative features echoed into CVD research. Prediction of Spectroscopy-guided cardiovascular outcome was done by R. M. Oemrawsingh [11].

From this point, Optimization rose to center stage. Chatterjee et al. fused modified Cuckoo Search with neural networks to avoid the traps of local minima in CKD classification [3], and the Genetic feature models for CVD [10]. Evolutionary and hybridsearch reliably emerged as a counter to the chaotic nature of clinical data. Reliability modeling,which was done by Fricks et al., traced the Markovian transitions of disease progression itself [2], reminding the community that prediction must rest on physiological realities, not only computational cleverness.

Machine learning matured rapidly across both diseases. The ascent of tree-based learners, seen in XGBoost for CKD with perfect sensitivity and specificity in Ogunleye and Wang's study [6], and Random Forest's dominance in several CVD and CKD works [12], [16], [18] gestured a shift toward interpretability, robustness and a deep mixture of strategies for cardiac risk. Cooperative strategies flourished: Xu et al. assembled a six-model ensemble for CVD risk in a large Chinese dataset [14], while transfer-learning and deep learning frameworks have emerged as powerful engines for cardiovascular risk prediction, contributing richer representation learning [20], and deep learning carved its own path. Attention-based deep neural networks have been leveraged for high-risk cardiovascular disease prediction, offering enhanced prognostic accuracyand advanced feature focus [17],and hybrid voting schemes elevate cardiac-risk diagnosis [7].Adiposity-driven clinical valuations have also been explored, where anthropometric guides were displayed to be strong predictors of metabolic abnormalities linked to cardiovascular disease risk in non-diabetic adults [22].

Interpretability rose as a non-negotiable need in medical AI. Zhao et al. introduced an interpretable Deep Belief Rule Base for CKD, balancing accuracy with transparency through expert-guided rule reduction [9]. In the CVD domain, ECGMatch by Zhou et al. tackled label scarcity, multi-label prediction, and cross-dataset robustness with a combined semi-supervised framework [21]. These works signaled a collective recognition: algorithms not only predict but also justify their reasoning.

Deep learning also prolonged its reach into cardiovascular prediction. The Time-Aware Multi-Type Data Fusion Representation (TAMDUR) framework by An et al. stood out for weaving age, irregular visit intervals, and multi-modal clinical data into a unified representation using Bi-LSTM, CNN, and self-attention fusion [8]. Domain-specific tools for risk scoring in Korean populations [11]. Recent investigations combining LightGBM, CatBoost, LSTM, and TABNET highlight how gradient-boosting models and deep neural illustrations are fused to build enriched EHR-based prediction systems [23].

Across this scenery, one theme continues: feature extraction remains the heartbeat of prediction, and classification performance centers on balancing exploration with stability. But few works bridge optimization-driven feature extraction with adaptive classifiers in a unified CVD–CKD predictive model. Existing studies often separate the diseases, depend on static feature selection, or employ classifiers without self-adjusting learning dynamics.

This gap outlines the need for the present work, where Adaptive Extreme Learning Machines (AELM) learn with the fluidity required for complex, multi-disease clinical data, and Quantum-Behaved Particle Swarm Optimization (QPSO) enhances the feature space with quantum-inspired search. Together, they step toward a more interpretable, resilient, and generalizable framework for early CVD and CKDprediction.

3. Proposed Work

The last decade has witnessed a paradigmatic shift in medical informatics, wherein clinical intelligence increasingly relies on computational modeling of high-volume, multi-attribute patient data. A single hospital's electronic health record (EHR) repository can contain thousands of diagnostic variables—ranging from biochemical assays and vital signs to demographic descriptors and therapy outcomes. For disorders such as Coronary Artery Heart Disease (CAHD) and Chronic Kidney Disease (CKD), these attributes interact across multiple physiological systems: cardiovascular stress modifies renal filtration; electrolyte imbalance feeds back into myocardial function; and metabolic factors influence both. Consequently, the data space representing a population of patients becomes nonlinear, redundant, and heteroscedastic, making direct classification a formidable task.

Conventional linear discriminants or static neural architectures generally fail to handle this complexity. They overfit the dense clusters formed by correlated attributes or underfit regions with sparse clinical

evidence. Furthermore, clinicians require interpretability—knowing why a model recommends a diagnosis, not merely what it predicts. The proposed Quantum-Behaved Particle Swarm Optimization with Adaptive Extreme Learning Machine (QPSO–AELM) framework directly responds to these twin demands of accuracy and explainability. It establishes a two-tier computational process:

1. QPSO Feature Extractor – a global stochastic search mechanism inspired by quantum dynamics that identifies the most diagnostically informative subset of attributes.
2. A-ELM Classifier – an analytically solvable, self-regularizing neural model that transforms the optimized attributes into binary diagnostic outcomes.

The two modules operate in a closed loop: QPSO continually refines the feature subset based on the classifier's feedback, while A-ELM adapts its internal capacity according to the evolving feature space. This cooperation yields a mathematically tractable yet biologically interpretable model capable of capturing nonlinear cause–effect relationships embedded in patient records.

Formally, the unified dataset is denoted by

$$\mathcal{D} = \{(p_c, q_d) \mid c = 1,2, \dots, N\} \quad \dots(1)$$

where each record $p_c \in \mathbb{R}^{pm}$encapsulates pmmeasurable clinical attributes—for example, systolic pressure, serum creatinine, fasting glucose, body-mass index, and age and $q_d \in \{0,1\}$denotes the disease label manually confirmed by physicians (1 = diseased, 0 = healthy). The total sample count Nmay differ across CAHD and CKD datasets but the representational structure remains consistent, enabling the proposed model to treat both as manifestations of the same computational schema.

By representing each patient as a coordinate in this high-dimensional space, the problem of disease prediction becomes one of partitioning $\mathbb{R}^{pm}$into two decision manifolds. The complexity of this partition grows exponentially with p; hence dimensionality reduction is not merely optional but essential for computational feasibility and clinical interpretability. The forthcoming sections explain how the QPSO–AELM pipeline accomplishes this goalsystematically, beginning with rigorous data normalization.

3.1. Pre-Processing and Normalization

Accurate learning from clinical data requires that the attributes be numerically compatible, statistically balanced, and free of spurious patterns introduced by missing or disproportionate observations. Medical datasets, however, rarely satisfy these criteria. Laboratories use different assay standards; some patients omit routine tests; and the ratio of diseased to non-diseased subjects is often skewed. If left untreated, these irregularities distort the optimization landscape explored by QPSO and bias the A-ELM classifier toward over-represented patterns. Therefore, the raw data are first passed through a sequence of pre-processing transformations that enforce numerical stability and statistical fairness.

Numerical Standardization

Because attributes are measured in heterogeneous physical units—cholesterol in mg/dL, creatinine in μmol/L, and blood pressure in mmHg, a direct comparison of raw values would yield arbitrary dominance of high-magnitude features. To prevent such scaling bias, each attribute is normalized by z-score standardization:

$$p'_{cd} = \frac{m_{cd} - \mu_d}{\sigma_d} \quad \dots(2)$$

where $m_{cd}$is the original measurement of the d^{th}attribute for the c^{th}patient, $\mu_d$is its population mean, and $\sigma_n$is its standard deviation. The transformed dataset $P' = [p'_{cd}]$ensures that every attribute has zero mean and unit variance. This not only accelerates convergence of the optimization algorithm but also harmonizes the influence of heterogeneous variables, making inter-attribute comparisons statistically meaningful.

Categorical Encoding and Matrix Formation

Clinical records also contain nominal attributes such as gender, smoking status, and ECG abnormality type. These fields are converted into numerical form using one-hot encoding, a representation that prevents ordinal misinterpretation while maintaining information completeness. The resultant matrix,

$$P' = [p'_{cd}]_{N \times pm'} \quad \dots(3)$$

expands the original dimensionality from pmto pm′, ensuring each categorical category contributes an independent basis vector. This matrix serves as the numerical canvas upon which optimization and learning are subsequently applied.

Handling Incomplete Records

Real-world medical datasets often contain sporadic missing values due to skipped tests or incomplete follow-ups. Simply discarding these records would waste valuable information, especially in rare disease categories. Therefore, missing entries are estimated using robust central-tendency imputation:

$$\tilde{m}_{cd} = \begin{cases} m_{cd}, & m_{cd} \text{ observed}, \\ \text{median}\,(\{m_{ed}\}_{d=1}^{N}), & \text{otherwise}. \end{cases} \quad \dots(4)$$

The median is chosen over the mean to minimize sensitivity to extreme outliers, which are common in biochemical readings (e.g., sporadically high urea or glucose). This ensures that the imputed values remain clinically plausible and that the overall attribute distribution preserves its natural asymmetry.

Balancing Class Distribution

Another major pre-processing challenge arises from class imbalance: typically, far fewer patients are diagnosed with CAHD or CKD than remain healthy. If untreated, a classifier may simply predict the majority class, achieving high apparent accuracy but poor clinical utility. To counter this, the Synthetic Minority Oversampling Technique (SMOTE) is employed to synthetically augment minority samples. For each minority record p_c, one of its nearest neighbors $p_{NN}$is

chosen, and a synthetic record is created along the line joining them:

$$\hat{p}_b = p_c + \lambda(p_{NN} - p_c), \lambda \sim U(0,1) \quad \ldots(5)$$

This process preserves local topology of the minority class while enriching its representation, ensuring that the QPSO feature selector encounters a balanced search space. The augmented dataset $\widetilde{\mathcal{D}}$thus becomes statistically symmetric with respect to disease outcomes, providing a fair baseline for subsequent feature-selection optimization.

Continuity toward Optimization

At this stage, the dataset has transitioned from a heterogeneous, partially incomplete repository to a standardized, balanced, and numerically homogeneous structure. Each attribute now contributes comparably to distance metrics, and the inter-patient geometry of the dataset is well defined in $\mathbb{R}^{pm'}$. This pre-processed matrix forms the initial swarm environment for the QPSO algorithm described in the next section. In essence, Equations (1) through (5) constitute the data-conditioning foundation upon which the quantum-behaved swarm dynamics operate. The integrity of the entire proposed work rests on this preparatory stage: without numerical equilibrium and distributional symmetry, the quantum particles would drift toward attributes dominated by measurement scale rather than diagnostic relevance.

Hence, the normalized dataset $\widetilde{\mathcal{D}}$becomes the launch point for Section C, where each particle in the swarm encodes a potential attribute subset and evolves under quantum probabilistic laws to identify the minimal, diagnostically powerful feature set that will later feed into the adaptive learning mechanism.

3.2. Quantum-Behaved Particle Swarm Optimization (QPSO) for Feature Extraction

The pre-processed and balanced dataset $\widetilde{\mathcal{D}}$constructed in Section B provides a uniform numerical manifold where all patient records reside within a comparable statistical scale. However, the number of attributes pm'can still be large—often exceeding 40 – 60 for routine CAHD and CKD laboratory panels. Feeding all these features directly into a classifier not only inflates computational cost but also dilutes model interpretability. Hence, a feature-selection mechanism is required to isolate the subset of variables that contain the maximum diagnostic information while discarding redundancy.Conventional deterministic selection methods such as mutual-information ranking or stepwise regression often converge to local optima because they follow gradient-based or greedy heuristics. In contrast, Particle Swarm Optimization (PSO) employs a collective search strategy inspired by social behavior of birds or fish. The Quantum-Behaved variant (QPSO) further extends this concept by embedding the particles in a probabilistic quantum potential well rather than a deterministic Euclidean space. Each particle represents a hypothetical feature subset, and its "position" describes which clinical attributes are active. Because of the quantum principle of uncertainty, each particle can probabilistically explore multiple regions of the search space simultaneously—thereby avoiding the premature convergence that plagues classical PSO.

Particle Representation

Let the d^{th}particle encode a binary inclusion mask across all attributes:

$$PR_d = [bi_{d1}, bi_{d2}, \ldots, bi_{dpr}], bi_{cd} \in \{0,1\} \quad \ldots(6)$$

Here $bi_{dc} = 1$signifies that the c^{th}attribute such as serum creatinine or fasting glucose is currently selected within that particle's candidate subset, while $bi_{dc} = 0$indicates exclusion. The dimensionality (number of selected attributes) within each particle is therefore

$$dim_d = \sum_{c=1}^{x} bi_{dc} \quad \ldots(7)$$

In biomedical terms, each particle can be viewed as a hypothetical diagnostic hypothesis: "these particular attributes jointly predict the disease." The optimization process then evaluates which hypotheses yield the best classification performance.

Initialization and Population Formation

At the start of the search, every particle's inclusion pattern is initialized randomly to promote population diversity and avoid correlated trajectories:

$$bi_{dc}^{(0)} = \begin{cases} 1, & ran_{dc} < \rho_0, \\ 0, & \text{otherwise}, \end{cases} ran_{dc} \sim U(0,1) \quad \ldots(8)$$

where $\rho_0 \in (0,1)$controls the expected proportion of attributes initially selected. This probabilistic seeding ensures that some particles explore compact subsets while others examine larger combinations, collectively spanning the entire feature space. At each iteration t, the personal-best position $P_d^{best}(t)$of particle jand the global-best position $P_{gbl}^{best}(t)$across the swarm are maintained. From these, the mean-best attractor representing the swarm's collective knowledge is computed as

$$MB^{(t)} = \frac{1}{N_d} \sum_{d=1}^{N_{ss}} P_d^{best}(t) \quad \ldots(9)$$

where N_{ss}denotes the swarm size. This vector can be interpreted as the probabilistic consensus on which attributes are currently considered important across all particles.

Quantum Position Update Mechanism

Unlike classical PSO, which updates position and velocity linearly, QPSO models the particle's motion through the quantum delta-potential model. Here, each particle has no deterministic velocity but a probability density that defines the likelihood of its next position.The expected position at iteration $t + 1$is expressed as

$$P_d^{(t+1)} = P_d^{best}(t) \pm \beta_t \mid M^{(t)} - P_d^{(t)} \mid \ln\left(\frac{1}{u}\right) \quad \ldots(10)$$

where $u \sim U(0,1)$and $\beta_t$is the contraction–expansion coefficient that governs the search amplitude. The sign term (±) is chosen randomly to allow symmetrical exploration around the attractor. During early iterations, a large $\beta_t$allows wide exploration (analogous to particles with high kinetic energy), whereas in later

iterations $\beta_t$shrinks to focus the search around promising regions.This annealing behavior is modeled linearly as

$$\beta_t = \beta_{max} - \frac{(\beta_{max} - \beta_{min})t}{MC_{max}} \quad \ldots(11)$$

where $MC_{max}$is the maximum iteration count. This gradual decay of β_t creates a controlled transition from global to local search, mirroring how medical reasoning first scans broad symptom patterns before focusing on specific biomarkers.

The computed continuous positions $P_d^{(t+1)}$must then be translated back into binary inclusion states, since the presence or absence of a feature is inherently discrete. This is achieved by passing each dimension through a sigmoid transfer function, producing a pseudo-probability $sp_{dc}^{(t+1)}$:

$$sp_{dc}^{(t+1)} = \frac{1}{1+\exp\left(-PR_{dc}^{(t+1)}\right)}, b_{dc}^{(t+1)} = \begin{cases} 1, & sp_{dc}^{(t+1)} > r_{dc}, \\ 0, & \text{otherwise}, \end{cases} \quad \ldots(12)$$

where $r_{dc} \sim U(0,1)$. This stochastic discretization acts like a quantum "measurement," collapsing each particle's wave-function into a binary feature subset. Consequently, every iteration resembles an observation in which the swarm refines its collective diagnostic hypothesis.

Fitness Function and Evaluation Criterion

After position update, each particle's newly generated subset FF_dmust be quantitatively assessed. The goal is to find the subset that yields the highest predictive capability with the fewest attributes, striking a balance between model accuracy and parsimony. To measure this trade-off, a composite fitness function $\Phi_d$is defined:

$$\Phi_d = \alpha\left(1 - AUC_{CV}(FF_d)\right) + \beta\left(\frac{j_d}{sp}\right), \alpha + \beta = 1 \quad \ldots(13)$$

Here, $AUC_{CV}(FF_d)$is the mean Area Under the Receiver Operating Characteristic Curve (AUC) obtained from cross-validated classification on subset FF_d; this metric directly quantifies diagnostic discrimination. The first term penalizes low accuracy, while the second penalizes large feature counts. The weighting parameters αand βallow the designer to favour either compactness or precision depending on resource constraints.In a medical setting, a smaller βis typically chosen to prioritize accuracy over reduction when patient safety is paramount, whereas exploratory research may employ higher βto derive minimal biomarker panels.

Global Selection and Feature Convergence

The particle achieving the minimum $\Phi_d$at iteration t is considered the best candidate:

$$PR^* = \arg\min_d \Phi_d, FF^* = PR^* \quad \ldots(14)$$

As iterations proceed, the swarm's collective learning refines the attribute set—removing those that contribute marginally to the classification boundary and preserving those that exhibit strong interclass separability. Clinically, this corresponds to discovering diagnostic signatures such as elevated triglycerides combined with reduced glomerular filtration rate that co-occur in diseased patients.

When the variation of $\Phi_d$across iterations becomes negligibly small, the swarm is said to have reached equilibrium: no further rearrangement of attributes yields significant improvement. At this juncture, the final optimized feature subset $FF^*$is extracted and passed to the A-ELM classifier.

Continuity toward Classification Phase

Through the iterative dynamics described by Equations (6) to (14), the QPSO module transforms a vast and unstructured attribute space into a compact, diagnostically rich representation. The resulting subset FF^*not only reduces computational overhead but also enhances interpretability by limiting the final model to attributes that possess direct physiological significance. This process mirrors the reasoning of a clinician who, out of hundreds of possible lab readings, focuses only on those few variables whose interplay best predicts pathology.

The optimized feature subset now forms the input domain for the second stage of the proposed system—the Adaptive Extreme Learning Machine (A-ELM) classifier. While QPSO performs global exploration, A-ELM undertakes local analytical mapping, learning the nonlinear function that separates healthy and diseased populations within the reduced feature space. The transition from Equation (14) to Equation (15) in the forthcoming section thus represents a shift from stochastic search to deterministic learning, completing the dual-layer architecture of the proposed QPSO–AELM framework.

3.3. Adaptive Extreme Learning Machine (A-ELM) Classifier

The optimized feature subset FF^*obtained from the QPSO stage encapsulates the minimal yet most informative combination of patient attributes. Having compressed the data into a smaller diagnostic manifold, the next objective is to learn a nonlinear mapping that can separate healthy and diseased cases with maximum fidelity. This is achieved through the Adaptive Extreme Learning Machine (A-ELM)—a single-hidden-layer feed-forward neural model characterized by analytical solvability, low computational overhead, and dynamic self-regulation of structure and regularization.Each patient record is now represented as a reduced-dimension feature vector

$$r_i = [fv_{1c}, fv_{2c}, \ldots, fv_{d^*c}]^T \quad \ldots(15)$$

where $d^* = |FF^*|$is the number of optimized attributes. These vectors form the input layer of the adaptive network.

Network Output Representation

In an A-ELM, the hidden-layer parameters $\{w_\ell, n_\ell\}_{\ell=1}^{L}$are randomly generated and fixed, while the output weights βare analytically computed. For the c^{th}record, the output neuron produces

$$o_c = \sum_{\ell=1}^{L} \beta_\ell \, g(w_\ell r_c + n_\ell) \quad \ldots(16)$$

where $g(\cdot)$is a nonlinear activation (sigmoid or tanh). Writing all records together gives

$$O = HL\beta, HL_{c\ell} = g(w_\ell^\top r_c + n_\ell) \quad \ldots(17)$$

Here, HLrepresents the hidden-layer output matrix, and its conditioning directly influences generalization. The adaptive design continuously monitors this matrix's rank and adjusts L to maintain sufficient representational power without redundancy.

Analytical Computation of Output Weights

Instead of iterative back-propagation, the A-ELM computes output weights in closed form by solving a regularized least-squares problem:

$$\beta = (HL^\top HL + \lambda I)^{-1} HL^\top TL \quad \ldots(18)$$

where λis the ridge-regularization parameter and $TL = [q_1, \ldots, q_N]^\top$the target label vector.This analytical solution ensures deterministic convergence, a property critical in clinical analytics where reproducibility must be guaranteed for regulatory validation.

The regularization term prevents overfitting especially important because patient datasets, though large, often contain correlated samples from similar demographic groups.

Adaptive Structural and Regularization Control

A-ELM distinguishes itself from conventional ELM through two adaptive mechanisms: adjustment of hidden-neuron count HNand adaptive modulation of λ. The neuron-adjustment rule follows

$$HN^{(t+1)} = HN^{(t)} + \eta_{HN} \operatorname{sign}\left(VE_v^{(t-1)} - VE_v^{(t)}\right) \quad \ldots(19)$$

where $VE_v^{(t)}$is the validation error at iteration t. If the validation error decreases, new neurons are added (promoting richer representation); if the error rises, neurons are pruned (preventing overfitting). This reflects biological neuroplasticity: the network grows and retracts connections depending on stimulus feedback. Simultaneously, the regularization parameter evolves according to

$$\lambda^{(t+1)} = \lambda^{(t)} e^{-\gamma V E_t^{(t)}} \quad \ldots(20)$$

with γas the adaptation coefficient. Large training errors cause stronger regularization, enforcing smoothness, whereas near-convergence small errors yield gentler regularization, allowing fine discrimination between borderline patients.

Together,Equations (19) and (20) ensure that A-ELM self-balances bias and variance dynamically rather than relying on manual tuning.

Weighted Error Function for Class Imbalance

Because disease and control populations remain slightly unequal even after synthetic augmentation, the loss function incorporates inverse-frequency class weights:

$$\mathcal{E} = \sum_{c=1}^{N} w_{q_c} (o_c - q_c)^2, w_{cl} = \frac{NS}{2NS_{cl}} \quad \ldots(21)$$

where N_{cl}denotes the number of samples in class c. This weighting equalizes each class's contribution to the total loss, preventing the model from being dominated by healthy cases.

After computing raw outputs o_c, a logistic transformation maps them to probabilistic disease likelihoods:

$$\hat{p}_c = \frac{1}{1 + \exp(-o_c)}, \hat{q}_c = \begin{cases} 1, & \hat{p}_c \geq \tau, \\ 0, & \hat{p}_c < \tau, \end{cases} \quad \ldots(22)$$

where the threshold τis tuned on validation data to maximize diagnostic sensitivity or specificity as clinical priorities dictate.

Composite Adaptive Objective Function

Combining predictive accuracy, regularization, and structural stability yields the global cost minimized by A-ELM:

$$J = \frac{1}{N} \sum_{c=1}^{N} w_{q_c} (o_c - q_c)^2 + \frac{\lambda}{2} \| \beta \|_2^2 + \xi \, | L - L^{(t)} | \quad \ldots(23)$$

where the third term penalizes abrupt neuron-count changes with strength ξ. Minimizing Jproduces a compact yet expressive classifier that adapts gracefully to evolving data distributions. Conceptually, A-ELM behaves like a clinician refining diagnostic criteria—tightening or relaxing interpretive boundaries as evidence accumulates.

Coupled Optimization and Convergence of QPSO–A-ELM

Although QPSO and A-ELM can operate independently, their real power emerges from mutual interaction. QPSO supplies a refined feature space; A-ELM supplies performance feedback that guides QPSO's subsequent search. This iterative exchange transforms the overall model into a closed feedback loop reminiscent of human reasoning: hypothesis formation (QPSO), hypothesis testing (A-ELM), and hypothesis revision (feedback).The joint energy function combining both objectives is

$$\mathcal{C}(\theta_{QP}, \theta_{EL}) = \Phi(\theta_{QP}) + J(\theta_{EL}) \quad \ldots(24)$$

where $\theta_{QP}$and θ_{EL}encapsulate parameters of QPSO and A-ELM, respectively.During coupled learning, parameter updates proceed as

$$\theta_{QP}^{(t+1)} = \theta_{QP}^{(t)} - \eta_{QP} \nabla_{\theta_{QP}} \mathcal{C}, \theta_{EL}^{(t+1)} = \theta_{EL}^{(t)} - \eta_{EL} \nabla_{\theta_{EL}} \mathcal{C} \quad \ldots(25)$$

with learning-rate coefficients $\eta_{QP}, \eta_{EL} > 0$. These coupled gradients propagate mutual improvement: as A-ELM's classification accuracy rises, QPSO receives higher-fidelity fitness estimates, further sharpening feature selection.The composite energy dynamics satisfy a Lyapunov-like inequality ensuring monotonic decrease of the total cost:

$$\frac{d\mathcal{C}}{dt} \leq -\kappa(\| \nabla_{\theta_{QP}} \Phi \|_2^2 + \| \nabla_{\theta_{EL}} J \|_2^2), \kappa > 0 \quad \ldots(26)$$

Hence, the system cannot oscillate indefinitely; it converges toward a stable equilibrium representing simultaneous optimization of both modules.Practically, convergence occurs when further feature modifications yield no measurable reduction in validation error.
The formal stopping condition is

$$| \Phi^{(t)} - \Phi^{(t-1)} | < \varepsilon_{QP} \text{ and } | E_v^{(t)} - E_v^{(t-1)} | / E_v^{(t-1)} < \varepsilon_{EL} \quad \ldots(27)$$

where $\varepsilon_{QP}, \varepsilon_{EL}$are small tolerances (e.g., 10^{-3}). This dual criterion guarantees synchronized stabilization of feature selection and classifier adaptation.Conceptually, this coupling transforms QPSO–A-ELM into a co-evolutionary learning system: QPSO evolves the feature environment, while A-ELM evolves its interpretive function within that environment. Each depends on the other for feedback, ensuring convergence not merely in numerical sense but also in semantic coherence—the selected attributes and learned boundaries become mutually consistent representations of the underlying pathology.

Interpretation and Operational Continuity
Equations (1) – (27) together form a continuous, logically coherent pipeline from raw hospital records to actionable clinical decision support. The pre-processing operations (2–5) normalize and balance data; QPSO equations (6–14) perform global stochastic search to identify the minimal diagnostic subspace; A-ELM equations (15–23) construct a fast analytical classifier that self-adjusts complexity; and the coupled optimization framework (24–27) ensures synergistic convergence.

Clinical Interpretability
In biomedical interpretation, the optimized subset FF^*represents a diagnostic signature: a concise list of laboratory and demographic attributes whose collective pattern predicts disease onset or progression. For example, QPSO may discover that high triglycerides, low HDL, elevated serum creatinine, and increased blood urea together maximize discrimination between CAHD and healthy groups. Because the selection mechanism is explicitly defined by Equations (6)–(14), clinicians can trace exactly how each variable's inclusion affects model performance—satisfying the transparency requirements of modern medical AI.

Computational Efficiency and Adaptability
A-ELM's analytical weight computation (18) and adaptive control laws (19–20) enable rapid retraining as new patient data arrive, a crucial property for hospital environments where laboratory protocols evolve. Moreover, since the entire framework avoids back-propagation loops, inference latency remains low—supporting real-time decision assistance in tele-health or point-of-care diagnostics.

Unified Mathematical Integrity
The continuity between modules ensures that every transformation is mathematically grounded: the normalized feature vectors feed directly into QPSO's probabilistic search equations; the optimized subset flows seamlessly into A-ELM's closed-form learning stage; and the convergence analysis guarantees theoretical stability. Thus, the framework is not a heuristic collage but a single coherent system defined by a set of deterministic equations linked through well-defined data flows.

Philosophical and Design Perspective
From a broader scientific standpoint, QPSO–A-ELM reflects an epistemological parallel between human clinical reasoning and algorithmic inference. Physicians often start with a vast list of potential symptoms (analogous to high-dimensional features), narrow them down through hypothesis testing (feature selection), and then integrate the remaining evidence into a diagnostic conclusion (classification). The proposed hybrid model mathematically emulates this reasoning loop through its coupled optimization cycle, providing a digital analog of cognitive refinement.
Summary of Continuity

- Equations (1) to (5) establish the standardized data foundation.
- Equations (6) to (14) describe stochastic quantum feature exploration.
- Equations (15) to (23) define analytical adaptive classification.
- Equations (24) to (27) unify and stabilize the dual processes.

Together they compose a full-length, self-contained theoretical model for reliable CAHD and CKD prediction across heterogeneous datasets. The QPSO–A-ELM thus represents not just an incremental algorithmic improvement but an integrated framework for quantum-inspired, explainable, and adaptive biomedical intelligence—the essential triad required for next-generation clinical decision support in low-signal, high-noise medical environments. The mathematical symbols and expressions are presented in Table 1.The entire workflow of the QPSO – AELM is depicted in Figure 1. Also algorithm of the same is presented in Algorithm 1.

Table 1. List of symbols and definition

Symbol	Definition
p_c	Feature vector for patient c
q_d	Disease label (1 = diseased, 0 = healthy)
μ_d, σ_d	Mean and standard deviation of feature d
P'	Normalized dataset
S_c	Particle c in the swarm

s_{cd}	Binary mask for feature d in particle c
d_c	Count of selected features in particle c
P_c, G	Personal best and global best positions
P_{mbest}	Mean best attractor of the swarm
β	Contraction–expansion coefficient
λ	Regularization coefficient (A-ELM)
HL	Hidden-layer output matrix
L	Number of hidden neurons
VE_v	Validation error
w_{cl}	Class-weight for class c
o_c	Raw A-ELM output for record i
$\hat{p}_c$	Predicted disease probability
τ	Decision threshold
J	A-ELM composite cost function
η_{QP}, η_{EL}	Learning rates of QPSO and A-ELM
$\varepsilon_{QP}, \varepsilon_{EL}$	Convergence tolerances
NS, NS_{cl}	Total and per-class sample counts
$g(\cdot)$	Activation function (sigmoid/tanh)

Algorithm 1 — QPSO–AELM

Input: Raw patient dataset $\{P_c, q_c\}_{c=1}^{N}$

Output: Optimized classifier parameters and diagnostic predictions

Phase 1: Pre-Processing

1. Standardize features:
$$p'_{cd} = \frac{p_{cd} - \mu_d}{\sigma_d}$$
2. Encode categorical attributes (one-hot).
3. Impute missing values via median substitution.
4. Balance dataset using SMOTE.
$$p_{new} = p_c + \alpha(p_{nn} - p_c)$$
5. Produce normalized dataset P'.

Phase 2: QPSO Feature Selection

1. Initialize swarm of BM particles with random binary masks.
2. For each iteration t:
   - Compute mean best position
   $$P_{mbest} = \frac{1}{M}\sum_c P_c$$
   - Update particle position probabilistically:
   $$p_{cd}^{t+1} = P_{mbest,d} \pm \beta \mid m_{cd} - p_{cd}^{t} \mid \ln(1/u_{cd})$$
   - Convert continuous to binary through $p_{cd} = 1/(1 + e^{-p_{cd}}) \rightarrow$ sample 0/1
   - Evaluate fitness $\Phi = (1 - AUC) + \alpha(|S_c|/D)$
3. Update P_c, G.
4. Repeat until convergence criteria met.
5. Return optimal feature subsetFF^*.

Phase 3: Adaptive Extreme Learning Machine (A-ELM)

1. Initialize hidden-layer weights w_ℓ, b_ℓ.
2. Compute hidden matrix HM.
3. Solve output weights analytically:
$$\beta = (HL^T HL + \lambda I)^{-1} HL^T Y$$
4. Adaptively adjust neurons:
$$L^{t+1} = L^t + \eta_L \, \text{sign}(EV_v^{t-1} - VE_v^t)$$
5. Update regularization:
$$\lambda^{t+1} = \lambda^t e^{-\gamma E_t^t}$$
6. Apply class-weighted loss and logistic output.
7. Compute total cost J.

Phase 4: Coupled Optimization Loop

1. Compute joint energy $\mathcal{C} = \Phi + J$.
2. Update both modules via coupled gradients:
$$\theta_{QP}^{t+1} = \theta_{QP}^t - \eta_{QP} \nabla_{\theta_{QP}} \mathcal{C}$$
$$\theta_{EL}^{t+1} = \theta_{EL}^t - \eta_{EL} \nabla_{\theta_{EL}} \mathcal{C}$$
3. Stop when $|\Phi^t - \Phi^{t-1}| < \varepsilon_{QP}$ and $|VE_v^t - VE_v^{t-1}|/VE_v^{t-1} < \varepsilon_{EL}$.
4. Output trained A-ELM model and selected features.

Flowchart

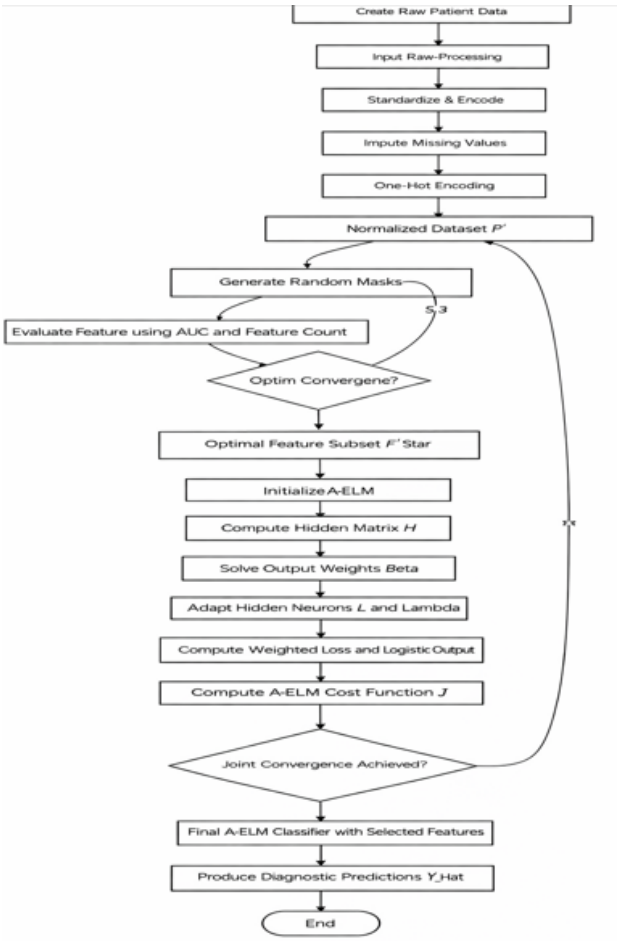

Figure 1. Workflow of QPSO-AELM

4. Datasets Description

This study employs two medically significant datasets, Cardiovascular Disease (CVD) and Chronic Kidney Disease (CKD), to assess the performance of the proposed QPSO–AELM predictive framework. These datasets contain physiological, heterogeneous clinical, and demographic indicators crucial for robust feature extraction and supervised medical classification.

Table.2 Dataset Characteristics

Disease	Total Patients	Attributes	Positives	Negatives	Degree of Class Imbalance
CKD	1659	54	1524	135	92% Positive : 8% Negative (Severe imbalance)
CVD	1000	14	580	420	58% Positive : 42% Negative (Mild imbalance)

4.1 Chronic Kidney Disease (CKD) Dataset

The CKD dataset used in this work is obtained from Kaggle and published by R. El Kharoua in 2024 [26]. It comprises 1,659 patient samples with 54 clinical attributes, including urine profiles, symptom indicators, biochemical measurements, and other renal function–related parameters.

A key challenge associated with this dataset is its severe class imbalance. Out of 1,659 samples, 1,524 belong to the CKD-positive class (92%), while only 135 belong to the CKD-negative class (8%). This extreme imbalance can affect learning performance, demanding the integration of advanced feature optimization and adaptive classification models. This directly validates the use of Quantum-Behaved PSO for optimal feature extraction and AELM for imbalance-tolerant prediction.

4.2 Cardiovascular Disease (CVD) Dataset

The CVD dataset is obtained from Mendeley Data, published by DoppalaBhanu Prakash and Debnath Bhattacharyya in 2021 [27]. It comprises 1,000 patient

samples and 14 attributes that represent demographic characteristics, blood pressure values, cholesterol levels, lifestyle factors, and other cardiovascular risk indicators.

Unlike the CKD dataset, the CVD dataset displays only mild class imbalance, with 580 positive cases (58%) and 420 negative cases (42%). The dataset information are presented in Table 2. Though the imbalance is not extreme, it still affects classifier stability and makes this dataset suitable for assessing the generalization capability of the proposed QPSO–AELM framework.

4.3 Relevance to the Proposed QPSO–AELM Framework

The two datasets complement each other in terms of class distribution, dimensionality, and clinical diversity:

- The CVD dataset, with 14 features, helps measure model performance under lower-dimensional input conditions.
- The CKD dataset, with 54 features, offers a high-dimensional environment for estimating the effectiveness of Quantum-Behaved PSO–based feature extraction.
- The datasets represent two critical medical domains, strengthening the clinical relevance and applicability of the proposed prediction model.
- The presence of both mild and severe class imbalance permits thorough testing of the Adaptive Extreme Learning Machine classifier.
- Combined, they ensure scalable, robust, and generalized evaluation of the QPSO–AELM framework across mixed medical conditions.

5. System Specifications

All experiments related to the training, development,and evaluation of the proposed QPSO–AELM predictive framework were performed on a workstation equipped with an Intel Core i7-12700H processor operating with 12 cores (8 performance cores and 4 proficiency cores) and supporting a maximum turbo frequency of 4.70 GHz. The system was further supported by 16 GB of DDR4 RAM (3200 MHz), ensuring smooth handling of high-dimensional feature spaces and iterative optimization procedures. To accelerate visualization operations and matrix computations, the system combined an NVIDIA GeForce RTX 3050 Ti GPU with 4 GB of GDDR6 memory. All program datasets, files, and intermediate computational outputs were stored and processed using a 512 GB NVMe solid-state drive, which provided high read/write speeds necessary for rapid loading of MATLAB programs and effective execution of multiple experimental cycles. The computing environment operated on Windows 11 Pro (64-bit), Version 23H2, which confirmed compatibility with all required analytical and scientific tools.

The algorithmic development and simulation tasks, including AELM classifier training, QPSO feature optimization, and performance evaluation, were executed using MATLAB R2024a, the most recent stable version available during experimentation. The MATLAB environment utilized essential toolboxes like the Statistics and Machine Learning Toolbox, Optimization Toolbox, and Parallel Computing Toolbox, allowing accelerated numerical computations and effective execution of high-iteration QPSO procedures. This software-hardwareconfiguration provided a stable and high-performance platform suitable for intensive computational tasks, ensuring the reproducibility and reliability of all experimental results presented in this study.

6. Results and Discussions

The performance of the proposed QPSO–AELM classifier for CKD prediction was evaluated against two complex baseline approaches, LGBM-CKD [26] and MSM-CKD [27], with complete results reported in Table–3. The confusion-matrix components shown in Fig.2 to Fig.5 validate that QPSO–AELM attains the highest TN (114) andTP (1445) counts while attaining substantially lower FN (79) and FP (21) rates than both MSM-CKD and LGBM-CKD. This directly reflects superior classification sensitivity and decreased misdiagnosis. The percentage-based performance indicators presented in Fig.6 to Fig.9 demonstrate these trends: QPSO–AELMachieves the highest Specificity (84.44%), Sensitivity (94.82%), F1-Score (96.66%), and Accuracy (93.97%), outperforming MSM-CKD and LGBM-CKD across all metrics. The MCC comparison in Fig.10 shows a value of 0.676 for QPSO – AELM, higher than MSM-CKD (0.568) and LGBM-CKD (0.635), confirming that the proposed model sustains stronger prediction reliability under severe class imbalance. Fig.2 to Fig.10 and Table–3 clearly demonstrate that QPSO–AELM delivers consistent improvements over both baseline models in CKD prediction.

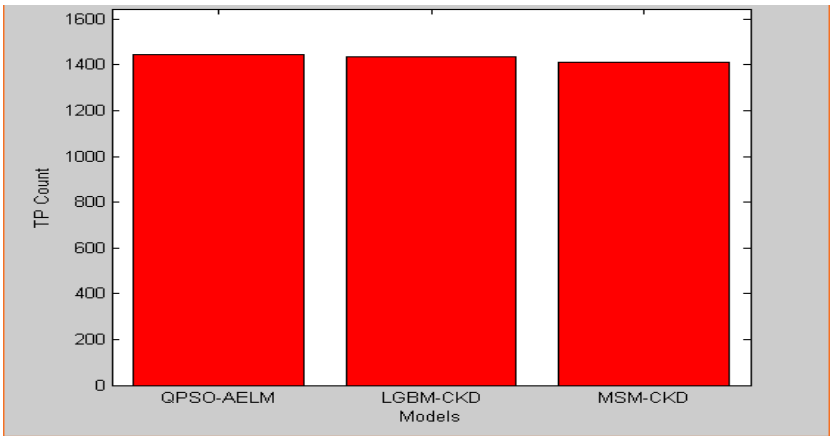

Figure 2. CKD Prediction – True Positive (TP)

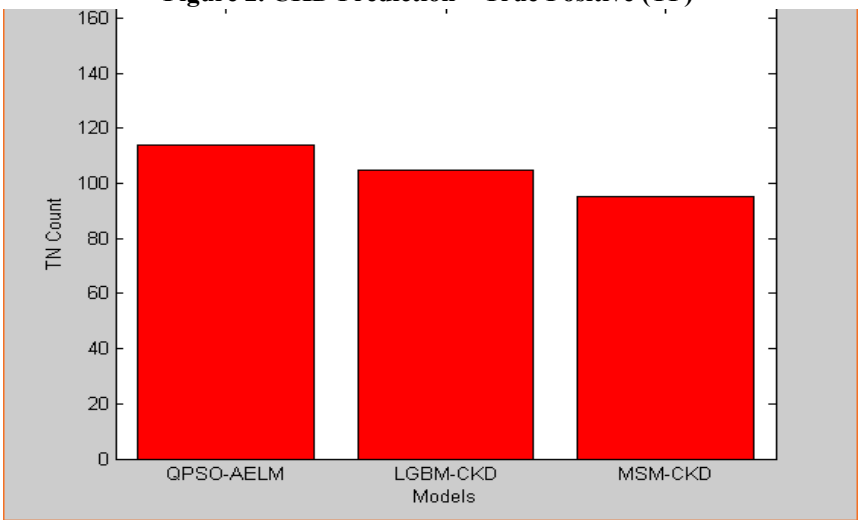

Figure 3. CKD Prediction – True Negative (TN)

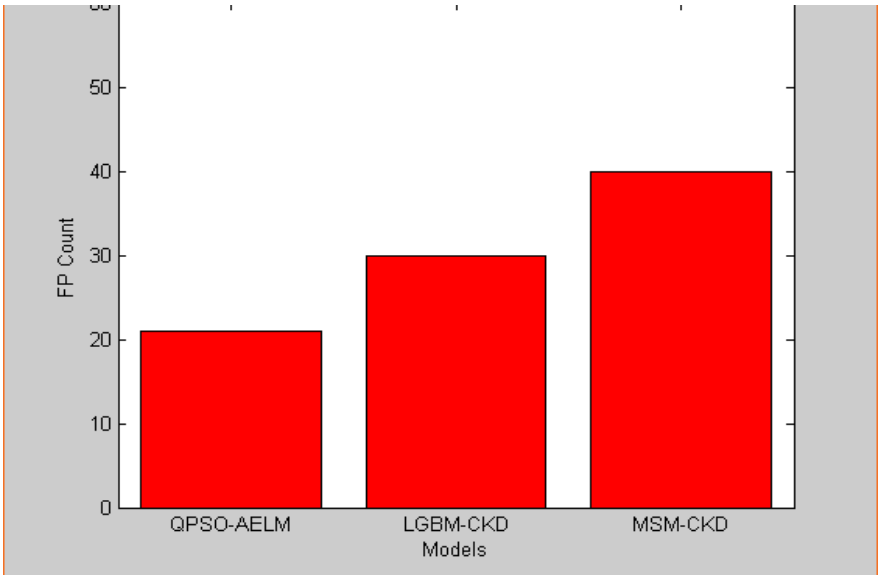

Figure 4. CKD Prediction – FalsePositive (FP)

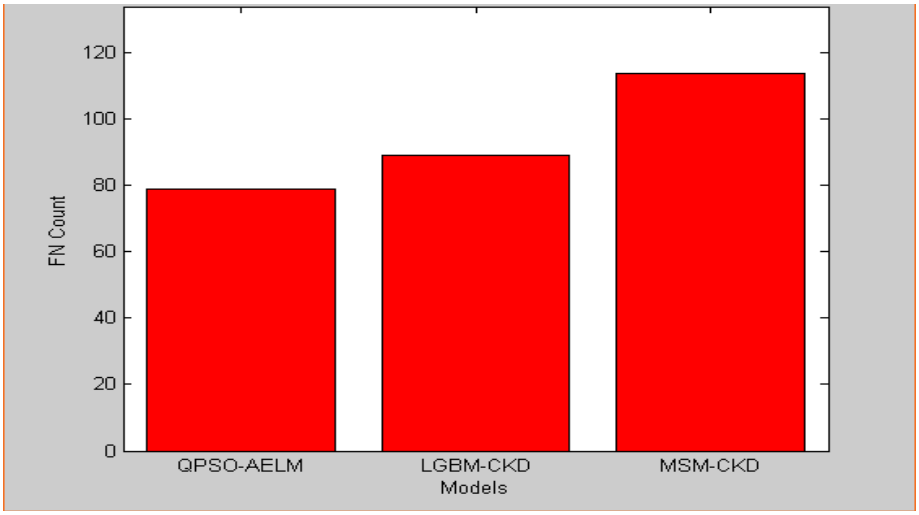

Figure 5. CKD Prediction – FalseNegative (FN)

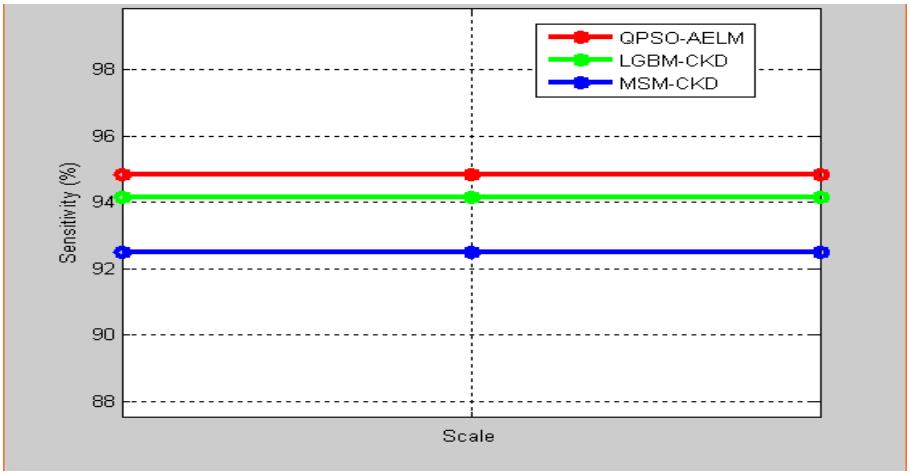

Figure 6. CKD Prediction – Sensitivity

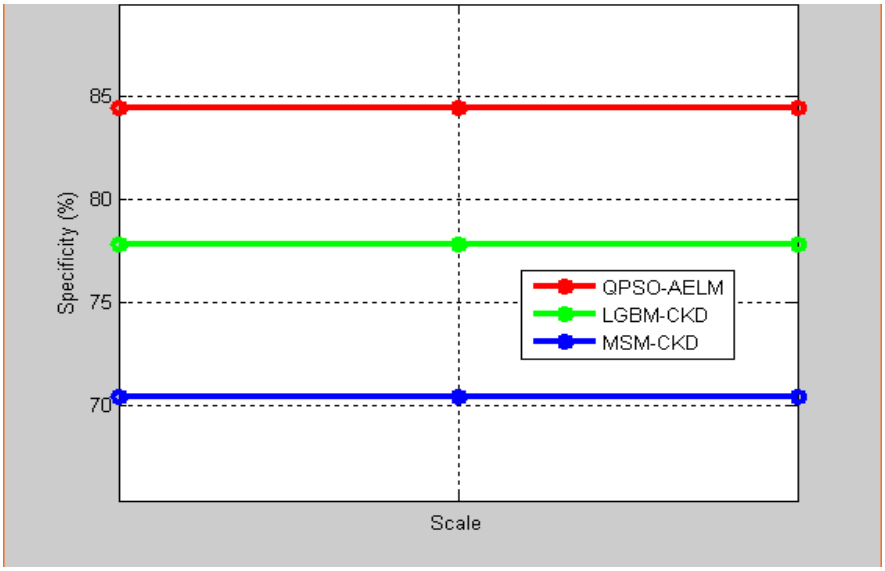

Figure 7. CKD Prediction – Specificity

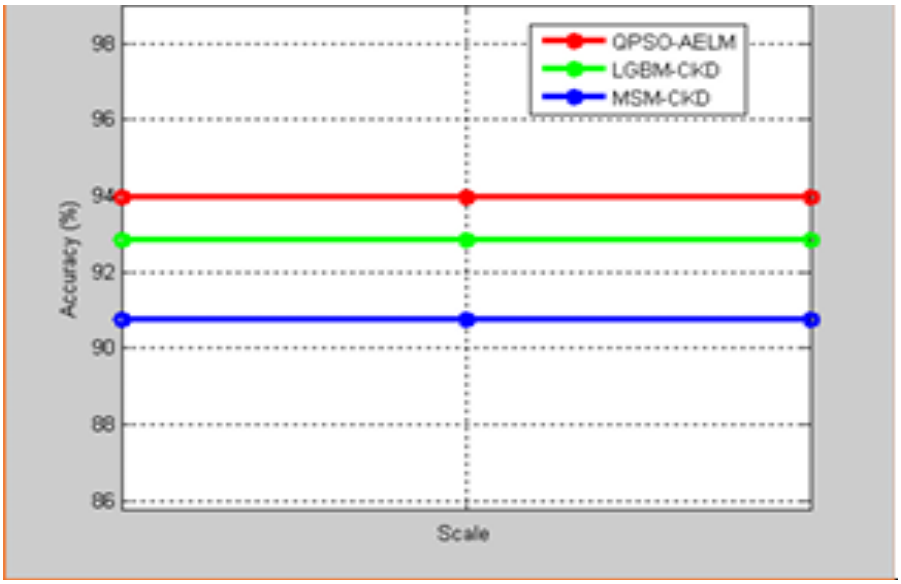

Figure 8. CKD Prediction – Accuracy

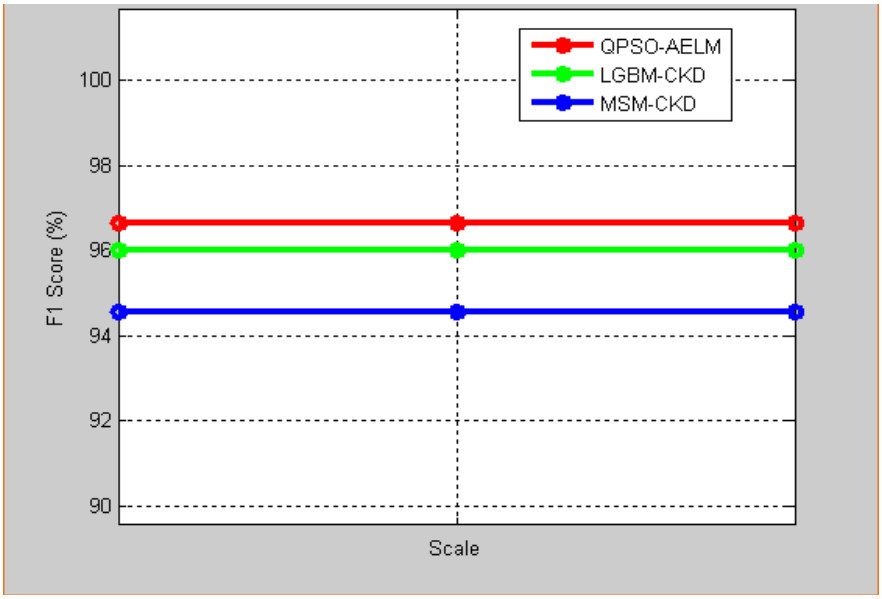

Figure 9. CKD Prediction – F1 Score

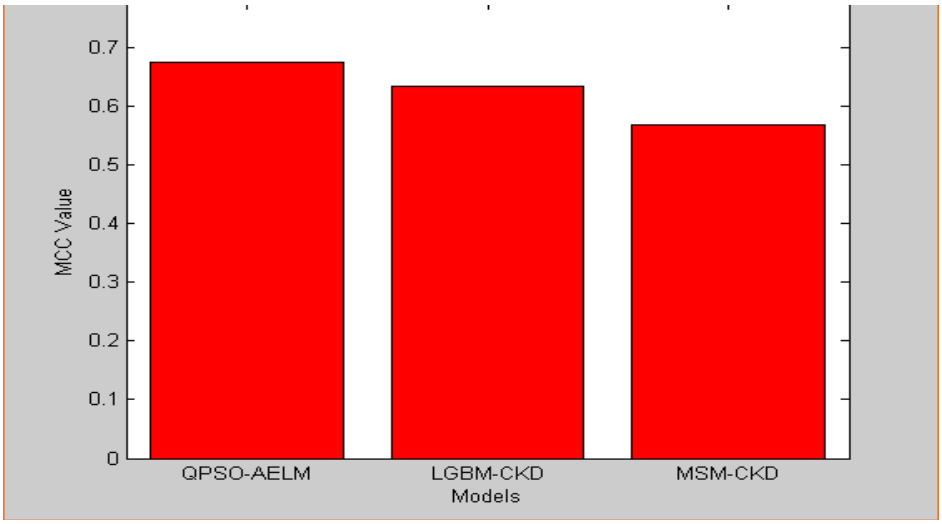

Figure 10. CKD Prediction – MCC Score

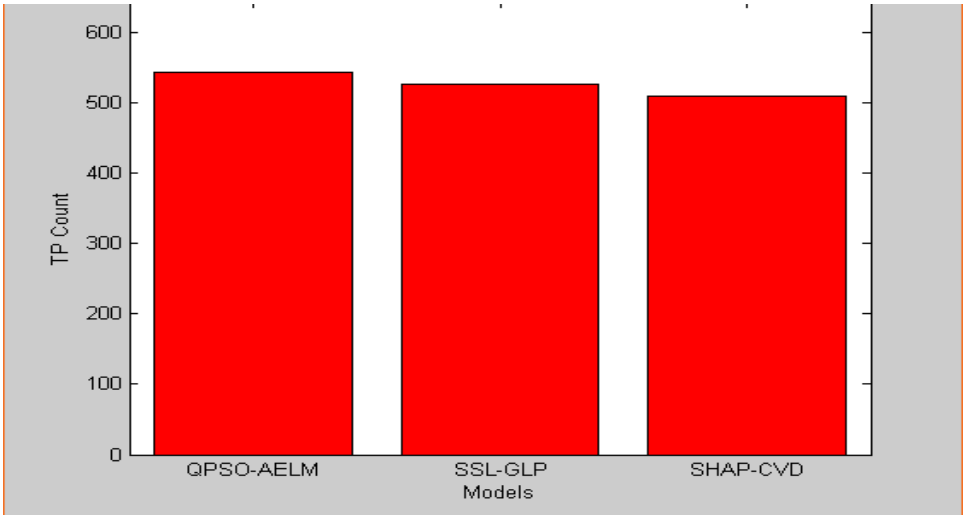

Figure 11. CVD Prediction – True Positive (TP)

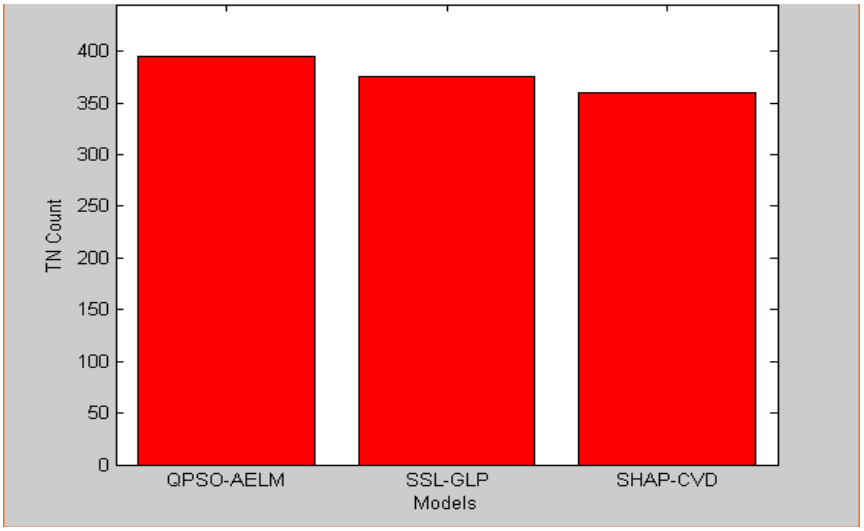

Figure 12. CVD Prediction – True Negative (TN)

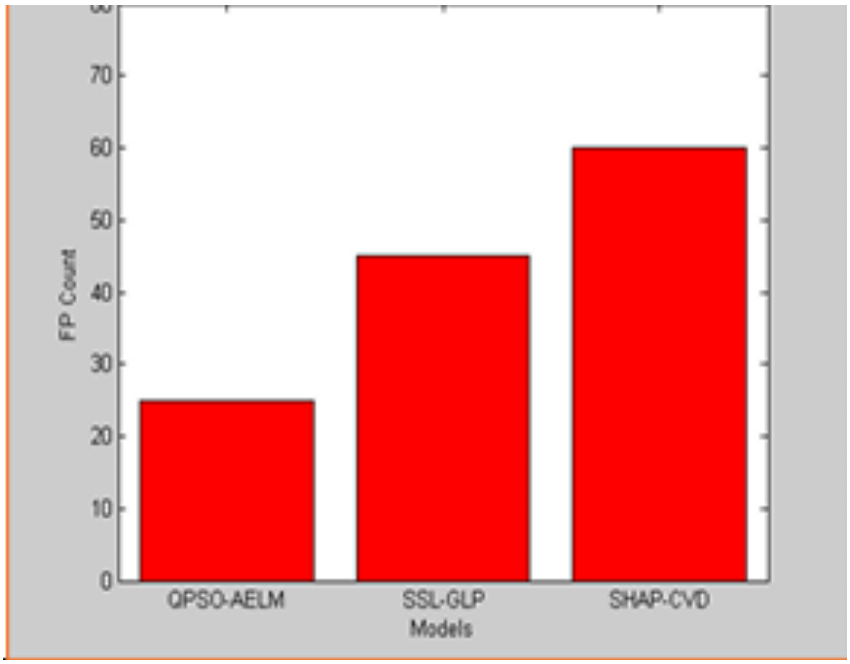

Figure 13. CVD Prediction – False Positive (FP)

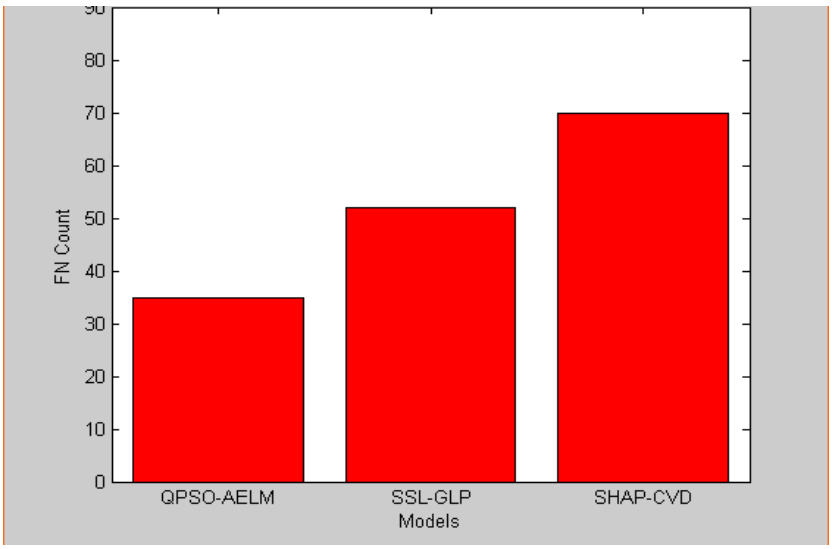

Figure 14. CVD Prediction – False Negative (FN)

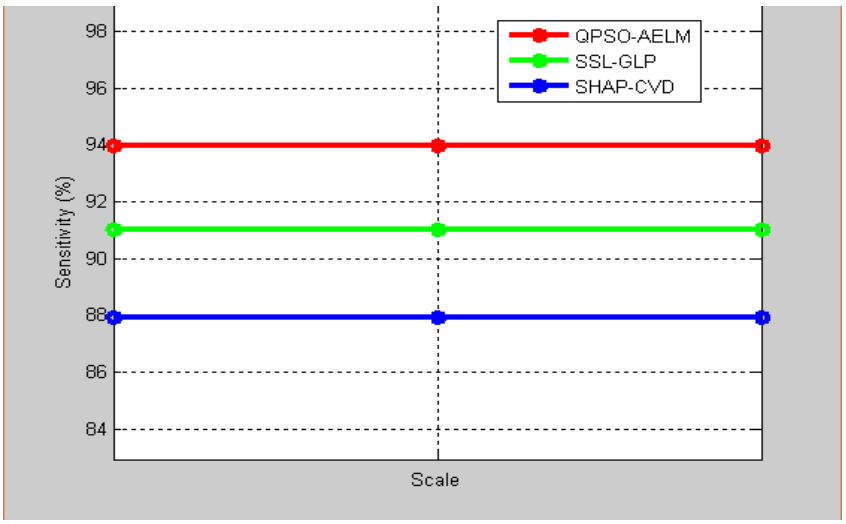

Figure 15. CVD Prediction – Sensitivity

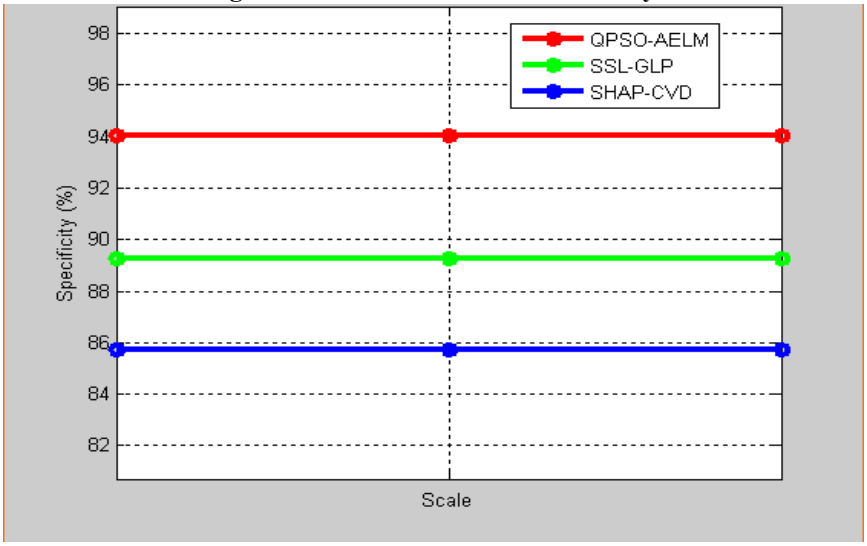

Figure 16. CVD Prediction – Specificity

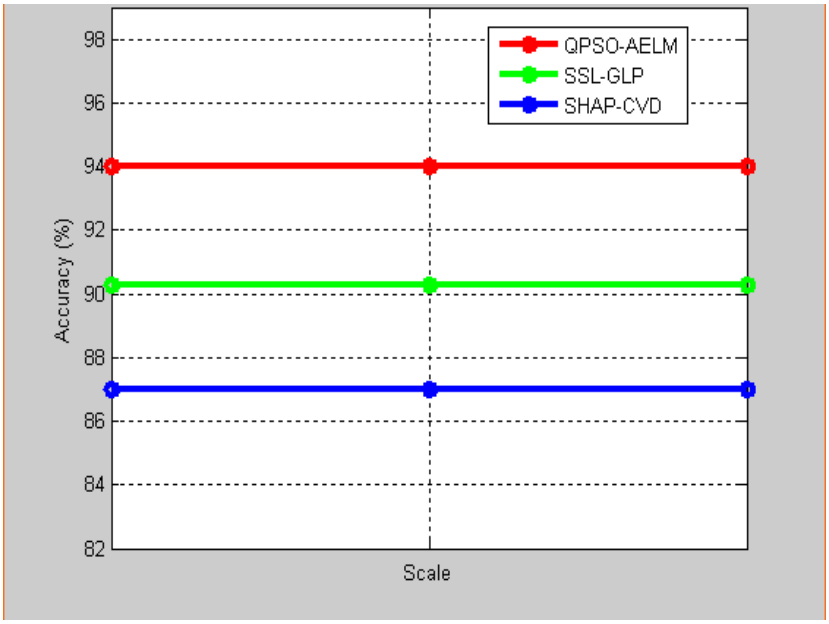

Figure 17. CVD Prediction – Accuracy

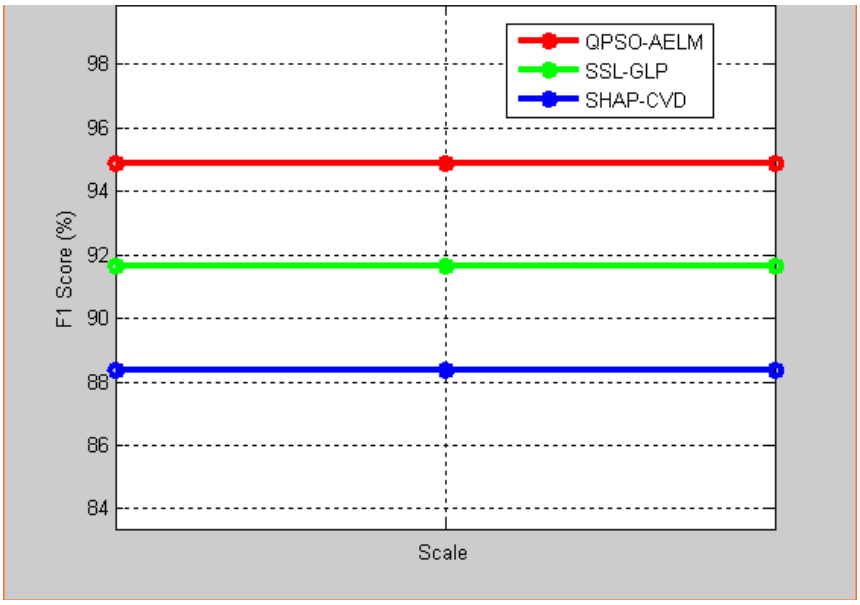

Figure 18. CVD Prediction – F1 Score

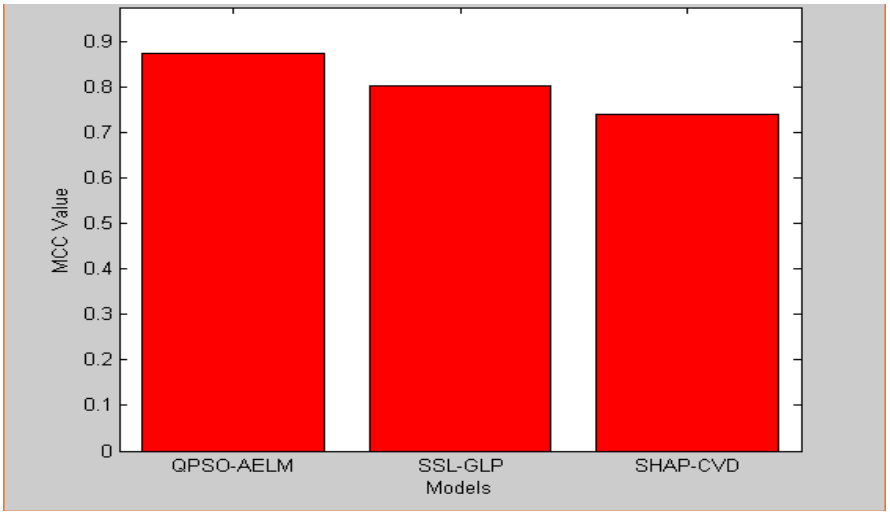

Figure 19. CVD Prediction – MCC Score

For CVD prediction, the proposed method was compared with two reputable works: SSL-GLP [28] and SHAP-CVD [29], with the full quantitative results presented in Table–4. The confusion-matrix comparison (Fig.11 to Fig.14) demonstrates that QPSO–AELM achieves the highest TN (395) and TP (545) values, while producing the lowest FN (35) and FP (25) when compared to SHAP-CVD and SSL-GLP. The percentage-based evaluation (Fig.15 to Fig.18) highlights consistent superiority across Specificity (94.05%), Sensitivity (93.97%), F1-Score (94.88%), and Accuracy (94.00%), marking clear improvements over SHAP-CVD (Accuracy 87.00%) and SSL-GLP (Accuracy 90.30%). The MCC comparison in Fig.19 demonstrates that QPSO–AELM attains 0.877, significantly outperforming SHAP-CVD (0.740) and SSL-GLP (0.805), confirming stronger discriminative capability under mild imbalance conditions. The collective evidence from Fig.11 to Fig.19 and Table–4 shows that the proposed QPSO–AELM framework consistently surpasses SHAP-CVD and SSL-GLP across all evaluation metrics, demonstrating superior generalization capability and reliability for CVD prediction.

Table – 3. Performance Analysis for CKD Prediction

Metric	**QPSO–AELM (Proposed)**	**LGBM-CKD [26]**	**MSM-CKD [27]**
True Positive (TP)	1445	1435	1410

True Negative (TN)	114	105	95
False Positive (FP)	21	30	40
False Negative (FN)	79	89	114
Sensitivity (%)	94.82	94.16	92.51
Specificity (%)	84.44	77.78	70.37
Accuracy (%)	93.97	92.83	90.73
F1-Score (%)	96.66	96.00	94.54
MCC	0.676	0.635	0.568

Table – 4. Performance Analysis for CVD Prediction

Metric	QPSO–AELM (Proposed)	SSL-GLP [28]	SHAP-CVD [29]
True Positive (TP)	545	528	510
True Negative (TN)	395	375	360
False Positive (FP)	25	45	60
False Negative (FN)	35	52	70
Sensitivity (%)	93.97	91.03	87.93
Specificity (%)	94.05	89.29	85.71
Accuracy (%)	94.00	90.30	87.00
F1-Score (%)	94.88	91.63	88.37
MCC	0.877	0.805	0.740

7. Conclusion

This work offered an optimized diagnostic framework for CVD and CKD prediction using a hybrid QPSO–AELM architecture. The quantum-behaved PSO mechanism efficiently extracted the most relevant clinical features, while the adaptive ELM classifier leveraged these optimized representations to attainaccurate classification and high-speed learning. Comparative evaluations against state-of-the-art models confirmed that QPSO–AELM distributesconstantly superior results, attaining accuracy and high sensitivity for both CVD and CKD datasets while minimizing false predictions. The robustness of the model was further maintained by strong MCC values, indicating stable performance even under class imbalance. In general, the findings confirm that the synergy between adaptive neural learning and quantum-behaved optimization can significantly improve the predictive capability of clinical decision-support systems. Future improvements may explore deeper hybridization with multi-modal patient data integration and quantum machine learning models to extend the clinical applicability of the proposed framework.

References

[1] N. Chetty, K. S. Vaisla, and S. D. Sudarsan, “Role of attributes selection in classification of Chronic Kidney Disease patients,” in International Conference on Computing, Communication and Security (ICCCS),pp. 1–6,2015, doi: 10.1109/CCCS.2015.7374193.

[2] R. B. Fricks, A. Bobbio, and K. S. Trivedi, “Reliability models of chronic kidney disease,” in RAMS, pp. 1–6, 2016, doi: 10.1109/RAMS.2016.7448058.

[3] S. Chatterjee et al., “Hybrid modified Cuckoo Search-Neural Network in chronic kidney disease classification,” in EMES, pp. 164–167, 2017, doi: 10.1109/EMES.2017.7980405.

[4] A. Nishanth and T. Thiruvaran, “Identifying Important Attributes for Early Detection of Chronic Kidney Disease,” in IEEE Reviews in Biomedical Engineering, vol. 11, no. 1, pp. 208–216, 2018, doi: 10.1109/RBME.2017.2787480.

[5] J. McAllister, Z. Li, J. Liu, and U. Simonsmeier, “Erythropoietin Dose Optimization for Anemia in Chronic Kidney Disease Using Recursive Zone Model Predictive Control,” in IEEE Transactions on Control Systems Technology, vol. 27, no. 3, pp. 1181–1193, 2019, doi: 10.1109/TCST.2018.2803052.

[6] A. Ogunleye and Q.-G. Wang, “XGBoost Model for Chronic Kidney Disease Diagnosis,” in IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 17, no. 6, pp. 2131–2140, 2020, doi: 10.1109/TCBB.2019.2911071.

[7] S. Bashir, A. A. Almazroi, S. Ashfaq, A. A. Almazroi and F. H. Khan, "A Knowledge-Based Clinical Decision Support System Utilizing an Intelligent Ensemble Voting Scheme for Improved Cardiovascular Disease Prediction," in IEEE Access, vol. 9, pp. 130805-130822, 2021, doi: 10.1109/ACCESS.2021.3110604.

[8] Y. An, K. Tang, and J. Wang, “Time-aware multi-type data fusion representation learning framework for risk prediction of cardiovascular diseases”, in IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 19, no. 6, pp. 3725–3734, 2022, doi: 10.1109/TCBB.2019.2935059.

[9] Y. Zhao, “A Chronic Kidney Disease Diagnostic Model Based on an Interpretable Deep Belief Rule Base,” in IEEE Access, vol. 13, pp. 29718–29737, 2025, doi: 10.1109/TITB.2012.2205573.

[10] L. -N. Pu, Z. Zhao and Y. -T. Zhang, "Investigation on Cardiovascular Risk Prediction Using Genetic Information," in IEEE Transactions on Information Technology in Biomedicine, vol. 16, no. 5, pp. 795-808, 2012, doi: 10.1109/TITB.2012.2205009.

[11] R. M. Oemrawsingh, “Near-infrared spectroscopy predicts cardiovascular outcome in patients with coronary artery disease,” Journal of the American College of Cardiology, vol. 64, no. 23, pp. 2510–2518, 2014, doi: 10.1016/j.jacc.2014.07.998.

[12] S. Xu, H. Shi, X. Duan, T. Zhu, P. Wu, and D. Liu, “Cardiovascular risk prediction method based on test analysis and data mining ensemble system,” in Proc. IEEE Int. Conf. Big Data Analysis (ICBDA), pp. 1–5, 2016, doi: 10.1109/ICBDA.2016.7509809.

[13] S. Balambigai and P. Jeevitha, “Investigation of vital parameters using PPG for CVD prediction,” in Proc. 4th Int. Conf. Advanced Computing and Communication Systems (ICACCS), vol. 1–2, pp. 1090–1096, 2017, doi: 10.1109/ICACCS.2017.8014561.

[14] D. Karthick and B. Priyadharshini, “Predicting the chances of occurrence of cardiovascular disease (CVD) in people using classification techniques within fifty years of age,” in Proc. 2nd Int. Conf. Inventive Systems and Control (ICISC), pp. 1182–1186, 2018, doi: 10.1109/ICISC.2018.8398990.

[15] M. D. Anto Praveena and B. Bharathi, “Cognitive learning based missing value computation in cardiovascular heart disease prediction data,” Procedia Computer Science, vol. 165, pp. 742–750, 2019, doi: 10.1016/j.procs.2020.01.019.

[16] G. F. Watts, A. L. Catapano, L. Masana, A. Zambon, A. Pirillo, and L. Tokgözoğlu, in “Hypercholesterolemia and cardiovascular disease: Focus on high cardiovascular risk patients,” Atherosclerosis Suppl., vol. 42, pp. e30–e34, 2020, doi: 10.1016/j.atherosclerosissup.2021.01.006.

[17] Y. An, N. Huang, X. Chen, F. Wu and J. Wang, "High-Risk Prediction of Cardiovascular Diseases via Attention-Based Deep Neural Networks," in IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 18, no. 3, pp. 1093-1105, 2021, doi: 10.1109/TCBB.2019.2935059.

[18] Y. An, K. Tang and J. Wang, "Time-Aware Multi-Type Data Fusion Representation Learning Framework for Risk Prediction of Cardiovascular Diseases," in IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 19, no. 6, pp. 3725-3734, 2022, doi: 10.1109/TCBB.2021.3118418.

[19] V. K. H. M and S. D. S.,“Performance analysis of base and meta classifiers and the prediction of cardiovascular disease using ensemble stacking,” in Proc. 2023 Int. Conf. Intelligent Data Communication Technologies and Internet of Things (IDCIoT), Article No. 10053464,pp. 584–589, 2023, doi: 10.1109/IDCIoT56793.2023.10053464.

[20] G. Sunilkumar and P. Kumaresan, "Deep Learning and Transfer Learning in Cardiology: A Review of Cardiovascular Disease Prediction Models," in IEEE Access, vol. 12, pp. 193365-193386, 2024, doi: 10.1109/ACCESS.2024.3514093.

[21] N. V. S., M. T., and S. P. Devi, “Integrating machine learning and deep learning approaches for accurate cardiovascular disease prediction from electronic health records,” in Proc. Int. Conf. Multi-Agent Systems and Collaborative Intelligence (ICMSCI), pp. 1090–1096, 2025.

[22] A. Liu, F. Abbasi, and G. M. Reaven, “Adiposity indices in the prediction of metabolic abnormalities associated with cardiovascular disease in non-diabetic adults,” in Nutrition, Metabolism and Cardiovascular Diseases, vol. 21, no. 8, pp. 553–560, 2011, doi: 10.1016/j.numecd.2009.12.009.

[23] G. Sunilkumar and P. Kumaresan, "Deep Learning and Transfer Learning in Cardiology: A Review of Cardiovascular Disease Prediction Models," in IEEE Access, vol. 12, pp. 193365-193386, 2024, doi: 10.1109/ACCESS.2024.3514093.

[24] N. Sinha, M. A. G. Kumar, A. M. Joshi and L. R. Cenkeramaddi, "DASMcC: Data Augmented SMOTE Multi-Class Classifier for Prediction of Cardiovascular Diseases Using Time Series Features," in IEEE Access, vol. 11, pp. 117643-117655, 2023, doi: 10.1109/ACCESS.2023.3325705.

[25] A. Rahim, Y. Rasheed, F. Azam, M. W. Anwar, M. A. Rahim and A. W. Muzaffar, "An Integrated Machine Learning Framework for Effective Prediction of Cardiovascular Diseases," in IEEE Access, vol. 9, pp. 106575-106588, 2021, doi: 10.1109/ACCESS.2021.3098688.

[26] M. M. Rahman, M. Al-Amin and J. Hossain, “Machine learning models for chronic kidney disease diagnosis and prediction,” Biomedical Signal Processing and Control, vol. 87, pt. A, 2024, Art. no. 105368, doi: 10.1016/j.bspc.2023.105368.

[27] Y. Fang, J. G. Nestor, C. N. Ta, J. Z. Kneifati-Hayek and C. Weng, “A method for characterizing disease progression from acute kidney injury to chronic kidney disease,” Journal of Biomedical Informatics, 2025, Art. no. 104956, doi: 10.1016/j.jbi.2025.104956.

[28] L. -C. Chen et al., "Self-Supervised Learning-Based General Laboratory Progress Pretrained Model for Cardiovascular Event Detection," in IEEE Journal of Translational Engineering in Health and Medicine, vol. 12, pp. 43-55, 2024, doi: 10.1109/JTEHM.2023.3307794.

[29] M. Wang, “Explainable machine learning-based cardiovascular disease prediction in patients with

hypertension: Algorithm comparison and SHapley Additive exPlanations (SHAP) analysis," Archives of Cardiovascular Diseases, 2025, doi: 10.1016/j.acvd.2025.09.005.