

ENHANCING GESTATIONAL DIABETES MELLITUS PREDICTION USING WHITE TIGER SWARM OPTIMIZATION-ENHANCED MULTILAYER PERCEPTRON (WTSO-MLP)

D.SHOBANA^{1,2}, V VINODHINI³

¹Research Scholar, Department of Computer Science, KPR College of Arts, Science and Research,
Coimbatore, India

²Assistant Professor, Department of Information Technology and Cognitive Systems, Sri Krishna Arts and
Science College, Coimbatore, India

³Associate Professor, Department of Information Technology, KPR College of Arts, Science and Research,
Coimbatore, India

: ^{1,2} shobanadl@gmail.com, ³vvprof133@gmail.com

ABSTRACT

Gestational Diabetes Mellitus (GDM) is a critical pregnancy-related complication that affects both maternal and neonatal health. Although commonly used for diagnosis, the Oral Glucose Tolerance Test (OGTT) is invasive, time-consuming, and fails to provide early detection. Inconsistent screening guidelines further complicate the identification of high-risk pregnancies, emphasizing the need for more accurate and timely predictive tools. This research develops a robust, non-invasive prediction model for GDM risk using White Tiger Swarm Optimization-enhanced Multilayer Perceptron (WTSO-MLP). The goal is to enhance early detection by integrating bio-inspired optimisation techniques to improve model performance while reducing dependency on invasive tests, such as OGTT. The WTSO-MLP model combines White Tiger Swarm Optimisation (WTSO) with Multilayer Perceptron (MLP) to optimise weight configurations, trained on a dataset of 3,525 instances that contain clinical and demographic data. Class imbalance has been addressed through adaptive techniques. Model performance has been evaluated using the Matthews Correlation Coefficient (MCC), Error Rate, Youden's Index, and Critical Success Index (CSI). This study contributes new knowledge by demonstrating how a bio-inspired optimization strategy can simultaneously refine neural network parameters and feature subsets using prospectively collected data, achieving superior accuracy, early detection capability, and adaptability across diverse clinical settings. The WTSO-MLP model outperformed traditional methods, achieving high performance in GDM prediction, especially for early-stage detection. The model demonstrated improved generalization, reduced misclassifications, and higher MCC scores, making it a reliable tool for clinicians. The WTSO-MLP model provides an innovative, efficient solution for early GDM risk prediction, improving diagnostic accuracy, generalization, and interpretability. It can seamlessly integrate into clinical workflows to enable early, non-invasive GDM assessments, ultimately enhancing maternal and fetal health outcomes.

Keywords: Healthcare, Diabetes, GDM, WTSO-MLP, Swarm Optimization

1. INTRODUCTION

Diabetes mellitus encompasses a group of metabolic disorders marked by elevated blood glucose levels, among which gestational diabetes mellitus (GDM) appears as a pregnancy-specific form. This condition emerges through impaired carbohydrate metabolism influenced by placental hormones [1]. GDM presents diagnostic challenges, as its onset is asymptomatic and confined to the gestational period. The condition usually develops between 24 and 28 weeks of pregnancy, though its

early pathophysiological effects begin much earlier. Global prevalence fluctuates widely due to variations in population genetics, clinical practices, and diagnostic criteria [2]. Such variability complicates risk assessment and standardization. With increasing maternal age, rising obesity rates, and sedentary lifestyles, GDM continues to rise globally. The lack of clear thresholds and universally accepted screening criteria has created uncertainty in identifying true risk groups [3]. Addressing this problem demands a framework that considers physiological, demographic, and clinical factors

while maintaining computational feasibility and fairness across diverse populations [4], [5].

GDM contributes substantially to adverse maternal and neonatal health outcomes, with direct effects on birth complications, delivery interventions, and postnatal care. These effects include excessive fetal growth, birth trauma, and maternal hypertensive disorders, which often require intensive monitoring and medical support [6]. Long-term metabolic consequences emerge in both the mother and child, including elevated risk of future type 2 diabetes and cardiovascular conditions. These risks lead to sustained healthcare engagement and recurring medical expenditure, especially in low-resource environments [7]. The socio-economic impact intensifies with indirect costs such as time away from work, specialized dietary planning, and limited access to prenatal services in remote or underserved regions. Diagnosis through laboratory testing increases this burden, particularly when unnecessary OGTT procedures are administered broadly without risk differentiation [8], [9]. Health disparities, care accessibility, and testing fatigue have made it essential to adopt targeted, early-stage diagnostic support that reduces avoidable intervention while prioritizing safety [10], [11].

Recent years have seen widespread interest in applying machine learning (ML) and deep learning (DL) to predict GDM from routine clinical records. These approaches rely on algorithmic models trained on historical data to classify high-risk cases based on medical and demographic inputs. While promising, such models often suffer from low sensitivity and reduced reliability across varying populations [12]. Using retrospective and imbalanced datasets, combined with limited diversity in training sources, introduces bias and lowers generalization capacity. Clinical practitioners hesitate to adopt these systems due to poor interpretability, unstable outputs, and limited control over decision boundaries [13]. Most models struggle with early-stage prediction, failing to capture subtle physiological indicators visible only in the initial trimester. Technical gaps in model design, absence of domain-aware learning constraints, and inability to quantify clinical confidence in outputs further restrict real-world use. These persistent deficiencies indicate a critical need to develop stable, explainable, context-aware frameworks trained on purpose-built, prospectively gathered data [14].

Bio-inspired optimization has emerged as a suitable enhancement to deep learning workflows, addressing training limitations through natural intelligence principles. Swarm-based techniques

imitate decision processes observed in nature to search for optimal model configurations under uncertain, non-linear conditions [15]. These strategies outperform traditional gradient-based learning in handling complex loss surfaces and feature dependencies. Inspired by adaptive behaviors in predatory environments, White Tiger Swarm Optimization (WTSO) incorporates stealthy exploration, responsive territory modulation, and controlled convergence [16]. These mechanisms contribute to a model's ability to refine weights adaptively without excessive drift. Multilayer Perceptrons (MLPs), known for their versatility in structured data learning, benefit from such guided optimization by avoiding overfitting and improving convergence stability. This adaptive potential becomes highly relevant in predicting GDM, where population heterogeneity, clinical variable interaction, and early physiological cues demand nuanced learning [17]. Combining biologically grounded optimization with an MLP architecture creates an informed pathway toward reliable GDM risk prediction with improved interpretability and clinical alignment.

GDM presents a critical public health challenge due to its rising global prevalence, potential for severe maternal–fetal complications, and strong link to future type 2 diabetes. Current screening practices, particularly OGTT, face limitations in accuracy, efficiency, and accessibility, leading to both over-testing and missed diagnoses. This research focuses on developing an advanced predictive model to enable early, reliable, and resource-efficient GDM risk identification, directly addressing these clinical and societal needs.

Bio-inspired optimization leverages patterns and strategies observed in nature to navigate complex solution spaces efficiently, adapting search behavior to dynamic problem landscapes [18]–[37]. By balancing wide-ranging exploration with focused refinement, it enables predictive models to reach optimal configurations that deliver both high accuracy and robust generalization, aligning closely with the desired research outcomes [38]–[57].

1.1 Problem Statement

GDM presents a persistent clinical challenge due to inconsistencies in screening guidelines, unreliable diagnostic thresholds, and variability in patient populations. Despite its substantial impact on maternal and neonatal outcomes, no universal consensus exists regarding whether all pregnant women should undergo oral glucose tolerance testing (OGTT) or only those in

defined risk groups. The OGTT poses challenges—invasive, time-consuming, costly, and often poorly tolerated—leading to suboptimal compliance. Existing risk-based models suffer from subjective definitions of risk factors and usually miss early-stage detection, particularly in the first trimester when intervention may be most effective. Moreover, variations in GDM prevalence (1–22%) across regions and datasets reflect the lack of standardized diagnostic practices and contribute to over- or under-diagnosis. Machine learning models proposed for GDM prediction have shown limited sensitivity and reliability, mainly due to their reliance on retrospective data, imbalanced class distributions, and inconsistent electronic health record (EHR) quality. These models often lack interpretability, which impedes clinical acceptance and regulatory approval. Real-world deployment further demands generalization across diverse populations and transparent reasoning in decision-making. These limitations collectively indicate a critical need for accurate, early, and generalizable predictive methods that reduce unnecessary testing while preserving clinical trust and diagnostic precision.

1.2 Motivation

GDM poses a serious public health challenge with far-reaching societal consequences. Its undetected progression not only jeopardizes maternal and neonatal health but also significantly increases the long-term risk of developing type 2 diabetes in both mother and child. Current screening methods, such as the oral glucose tolerance test (OGTT), are invasive, resource-intensive, and often inaccessible in rural or underserved regions, contributing to disparities in prenatal care. Inaccurate or delayed diagnosis leads to unnecessary interventions for some and missed preventive opportunities for others. The healthcare system bears the cumulative economic burden through increased hospitalizations, long-term diabetes care, and associated complications. Early and equitable identification of at-risk pregnancies would allow timely interventions, reduce complications, and optimize healthcare resource allocation. Societal well-being can be advanced through predictive tools that integrate with digital health systems, enabling cost-effective, scalable, and personalized risk assessment. Empowering clinicians and communities with such capabilities fosters preventive care, promotes health equity, and mitigates the growing intergenerational impact of diabetes across populations.

1.3 Objective

This research aims to develop an early, adaptive, and interpretable predictive model named WTSO-MLP (White Tiger Swarm Optimization-enhanced Multilayer Perceptron) for identifying individuals at risk of GDM before clinical manifestation. This work aims to overcome the core challenges in current GDM screening practices, such as dependence on invasive OGTT, ambiguous diagnostic thresholds, and variable classification accuracy in traditional machine learning approaches. The WTSO-MLP model integrates nature-inspired swarm intelligence with deep learning to dynamically refine neural weights, maintain prediction stability, and improve diagnostic sensitivity across diverse populations. Special emphasis has been placed on achieving high classification accuracy and enhancing interpretability using model-explainable outputs to facilitate clinical trust. The model includes an embedded strategy to regulate learning dynamics through a circadian rhythm-based learning rate and applies a multi-strategy optimization mechanism to escape local minima. The proposed framework is intended for practical integration into clinical decision-support environments, enabling personalized risk assessment, reduced diagnostic burden, and timely intervention, thereby promoting scalable, equitable, and preventive prenatal healthcare.

2. LITERATURE REVIEW

“Wearable TFT Insight” [58] integrates CGM streams with static user data into a Temporal Fusion Transformer (TFT) model that applies self-attention to emphasize recent trends. Static features like demographics are embedded into time-series input. After cloud training, the model is compressed and deployed onto a system-on-chip. The edge device processes glucose predictions in real-time using sensor input, temporal patterns, and compact computation, enabling personalized and low-power glucose forecasting without cloud reliance. “Three-Track Risk Grid” [59] combines regression-based statistical formulas, classical machine learning classifiers, and deep neural architectures to model diabetes progression. Regression maps biomarkers to outcomes, ML classifiers like SVM and Random Forests rank predictive features, and deep networks handle non-linear patient history. Data preparation includes normalization and missing data handling. Each model is validated using AUC and recall metrics. “Feature-to-Controller Bridge” [60] links glucose trends, insulin intake, and carb logs with rule-based, regression, and neural models using

constraint-aware learning. Data is preprocessed into structured sequences and routed into multiple predictive layers—clinical safety zones, like hypo/hyperglycemia thresholds, guide training targets. Output predictions are evaluated using real-world safety metrics and explained using SHAP/LIME. Feedback loops adjust insulin forecasts dynamically.

“Pathway Tree” [61] uses random forest survival models to trace transitions between immune states by capturing autoantibody timelines. Patient data is segmented into event intervals and encoded with biomarkers and history. Two forests separately forecast early immune conversion and clinical onset, using time-to-event learning. Each tree estimates transition likelihood by statistical splits. Importance scores rank features influencing disease risk. “HER-Aware Attention Net” [62] reformats inconsistent health records into ordered multi-feature maps capturing local and long-term signal patterns. Temporal derivatives and statistical summaries enhance variability tracking. Dual attention mechanisms highlight essential features and time frames. Time-aware embeddings handle missingness while medical ratings add clinical context. A classifier fuses all levels to output risk probabilities. “Stack-Guided Fusion Core” [63] merges CNNs, RNNs, and DNNs into a meta-learning stack where each base learner captures distinct clinical patterns. Raw medical data is scaled and fed to the learners, whose outputs are aggregated through a gradient-based meta-learner. Attention and calibration modules guide interpretability and confidence scoring. Final predictions emerge from a harmonized view of structural, sequential, and relational features.

“Demographic-Stratified Voting Grid” [64] segments the dataset by gender and age, fitting separate models to each group using scaled clinical inputs. Each subgroup uses tailored encoders and trains multiple learners, such as SVM, RF, and boosting. Predictions are aggregated with weighted voting based on subgroup performance. A calibration layer aligns subgroup outputs into a unified prediction. “Prior-Infused Bayesian Filter” [65] fuses CGM data with domain-informed priors using a Bayesian filtering pipeline. Glucose patterns, meal intake, and insulin history are observed features, while expert physiological expectations guide state transitions. Gaussian Process Regression models noise, and Kalman filtering smoothens outputs. Bayes’ rule updates posterior predictions with each time step, reflecting learned dynamics and clinical bounds. “Hierarchical Layer Fusion” [66] combines deep neural feature extraction with ensemble tree classifiers. Initial layers include

bidirectional LSTM and dense networks that learn sequence- and context-sensitive health patterns. These embeddings feed into classifiers like Random Forest and Gradient Boosting, optimized via Bayesian tuning. Softmax temperature scaling adjusts prediction certainty. SHAP values and class activation overlays identify dominant biomarkers.

“Confidence-Calibrated Meta Loop” [67] blends MAML-based meta-learning with evidential deep networks for personalized glucose prediction. Historical data per subject is parsed into learning episodes, from which a generalized model is fine-tuned with minimal steps. Evidential layers yield belief intervals alongside forecasts, while a critic loss penalizes overconfident errors. “Explanation Dual Lens” [68] compares LIME and SHAP interpretations over models predicting diabetes from structured clinical data. LIME builds local surrogate models through input perturbation to determine impactful features per instance. SHAP computes global feature contributions using Shapley values derived from all feature subsets. Tree SHAP accelerates evaluations for tree-based models. “Subspace Fusion Select” [69] integrates radiomic and clinical features using multi-view subspace clustering for diabetes risk after pancreatectomy. 3D scan-derived wavelet features and structured patient data are clustered into latent groups, identifying cross-view correlations. MSCUFS selects high-relevance features within and across views. Selected features are fused and passed to a classifier like SVM. Imaging and clinical metrics reinforce one another through clustering-guided fusion, modeling post-surgical diabetes onset with improved consistency.

“Edge-Attentive IoMT Loop” [70] enables wearable glucose forecasting through real-time deep learning on edge devices. Glucose readings are streamed into an attention-equipped recurrent network, prioritizing recent influential data points. The system operates on-device to minimize latency and compute offloading. A connected mobile app provides alerts and visualization. Model updates occur periodically in the cloud. “PA-BGL Fusion Layers” [71] incorporates heart rate and movement data with glucose trends using three fusion strategies. Signal-level fusion directly aligns sensor inputs, feature-level fusion derives statistical activity metrics, and decision-level fusion averages outputs from separate models. The system learns physiological reactivity from short bursts and prolonged behaviors. “Sensor-Stacked Fusion Grid” [72] aggregates time-series data from glucose, ECG, motion, and respiration sensors. Each signal undergoes feature extraction, producing structured

vectors. Sensor combinations are evaluated with XGBoost, determining which sensor sets yield optimal predictive accuracy. Fusion prioritizes glucose, ECG, and motion data. Redundant sensors are discarded.

“Adaptive Harmony-Driven Hybrid Stacking (AHDHS)” [73] explores a dual-optimized ensemble model suited for early gestational diabetes prediction. Unlike traditional stacking, which treats features and learners as static, AHDHS uses harmony search to select both dynamically. It eliminates redundant variables, such as correlated hormone indicators, while optimizing classifier synergy. The result is a leaner ensemble that responds well to limited, early-trimester datasets. This structure improves detection precision during early check-ups when symptom patterns are weak or variable, reducing dependence on late-stage lab-based diagnostics. “Outlier Detection with Deep Stacked Autoencoder (OD-DSAE)” [74] presents a robust mechanism for early GDM screening by integrating hierarchical anomaly detection with deep representation learning. Due to noisy data and unfiltered extremes, traditional clinical models often miss atypical cases. OD-DSAE removes such irregularities through mutual information-based clustering and learns discriminative representations via autoencoders. Unlike shallow predictors, it preserves patient-specific glucose dynamics while generalizing across screening windows. This hybrid ensures reduced false positives during early gestation, a critical phase for initiating dietary or insulin intervention.

2.1. Problem Hypothesis

It is hypothesized that the limitations of current GDM screening such as the invasiveness of OGTT, inconsistent diagnostic thresholds, and low sensitivity for early detection can be overcome by integrating a bio-inspired optimization strategy with a neural network classifier. By combining White Tiger Swarm Optimization (WTSO) with a Multilayer Perceptron (MLP), it is possible to simultaneously optimize model parameters and select the most relevant features from high-quality, prospectively collected clinical data. This approach is expected to improve predictive accuracy, enhance early-stage detection, reduce unnecessary testing, and adapt effectively across diverse patient populations.

3. PROPOSED METHODOLOGY

The White Tiger Swarm Optimization-enhanced Multilayer Perceptron (WTSO-MLP) model has been developed to predict Gestational

Diabetes Mellitus (GDM) using clinical and demographic data. This research integrates White Tiger Swarm Optimization (WTSO) with Multilayer Perceptron (MLP) to improve prediction accuracy while reducing reliance on invasive tests like OGTT. The model is trained and validated using a dataset of 3,525 instances, collected prospectively from 2019 to 2021. This study aims to identify individuals at high risk for GDM early, preventing unnecessary testing and improving timely interventions. The model focuses on providing an interpretable and computationally efficient solution that can be integrated into clinical decision support systems, helping optimize healthcare resources and improving maternal health outcomes.

3.1 Initialization

The initialization phase in WTSO-MLP defines the neural network architecture and encodes the search space for weight optimization. This phase establishes input feature representation, assigns initial weight values, and structures the swarm-based optimization framework inspired by white tiger behavioral mechanisms. Proper initialization enhances training stability, convergence speed, and search efficiency in GDM prediction.

A multilayer perceptron designed for GDM prediction consists of input, hidden, and output layers, where each neuron represents a clinical feature or activation unit. The initialization phase ensures that the feature representations capture non-linear relationships between biological and lifestyle factors affecting GDM risk.

$$X = \{x_1, x_2, \dots, x_n\},$$

$$x_i \in R, \quad i = 1, 2, \dots, n \quad (1)$$

$$W^l = \{w_{ij}^l \mid i \in [1, m], j \in [1, n]\} \quad (2)$$

The feature set X consists of n clinical parameters such as glucose level, BMI, insulin resistance, and blood pressure, while W^l represents the weight matrix in layer l , where m and n denote neuron counts in successive layers. Proper initialization of W^l is essential for stable gradient propagation and practical training.

Each white tiger agent in WTSO represents a unique candidate weight configuration that undergoes adaptive optimization. The initialization step assigns each agent a randomly generated weight set, ensuring diverse exploration in the search space.

$$\theta_i^{(0)} = \alpha \cdot N(0, \sigma^2) \quad (3)$$

$$P_{init} = \{\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_N^{(0)}\} \quad (4)$$

The initial weight parameter $\theta_i^{(0)}$ for each agent i , a Gaussian distribution $N(0, \sigma^2)$, where σ^2 is the variance control factor. The population P_{init} consists of N white tiger agents, each encoding a unique weight set. This design ensures search diversity, preventing premature convergence.

Inspired by white tigers' adaptive hunting strategies, the search region dynamically expands or contracts depending on convergence rates. This prevents excessive exploitation in the suboptimal areas and allows compelling weight exploration.

$$S_i^{(0)} = S_{min} + (S_{max} - S_{min}) \cdot e^{-\lambda t} \quad (5)$$

$$\Delta W_i^{(0)} = \eta \cdot (W_{max} - W_{min}) \cdot \xi \quad (6)$$

The search boundary function $S_i^{(0)}$ expands or contracts based on a decay factor λ and iteration step t . The perturbation function $\Delta W_i^{(0)}$ introduces controlled randomness, where η is an adaptive factor and ξ follows a normal distribution to introduce stochastic variations.

Bias parameters modify neuron activation thresholds, allowing flexibility in decision boundary shifts. An adaptive thresholding method inspired by tiger predation strategies has been applied to fine-tune bias values.

$$b_i^{(0)} = \gamma \cdot \tanh(\delta_i) \quad (7)$$

$$T_{threshold} = \frac{1}{N} \sum_{i=1}^N |\theta_i^{(0)}| \quad (8)$$

The initial bias vector $b_i^{(0)}$ follows a hyperbolic tangent transformation, where γ is a scaling factor, and δ_i is a randomized initialization term. The adaptive threshold function $T_{threshold}$ ensures stable gradient flow by normalizing initial weight magnitudes.

3.2 Swarm Generation

The swarm generation phase in WTSO-MLP has established an adaptive population-based search framework, where each white tiger agent represents a potential solution for neural network weight optimization. The initialization has ensured diversity in the search space by distributing agents across multiple regions, facilitating exploration and

exploitation during training. The designed swarm structure enhances the model's ability to detect non-linear patterns in GDM prediction, improving classification accuracy and computational efficiency.

The swarm consists of agents, each encoding a distinct set of weight parameters for the neural network. These agents navigate the search space, evaluating loss functions and adjusting positions based on an adaptive search mechanism. The population distribution has been structured to maintain diversity, ensuring efficient convergence to optimal weight configurations.

$$P = \{A_1, A_2, \dots, A_N\}, \quad A_i = \{\theta_{i1}, \theta_{i2}, \dots, \theta_{im}\} \quad (9)$$

$$M = \sum_{i=1}^N A_i \cdot \Psi(A_i) \quad (10)$$

The swarm set P comprises N agents, and each agent A_i storing a unique set of weight values θ_{ij} , where m represents the total parameter count. The adaptive weight modulation matrix M has been computed using the fitness-based modulation function $\Psi(A_i)$, ensuring an intelligent distribution of agents across the search space.

The placement of agents within the search space has been distributed dynamically, ensuring balanced coverage across different weight configurations. This structure allows the swarm to converge efficiently, without stagnating in suboptimal solutions.

$$S_i^{(0)} = \Omega_i \cdot \left(\frac{\Gamma_{max} - \Gamma_{min}}{N} \right) + \Gamma_{min} \quad (11)$$

$$V_i = \Lambda \cdot (S_i^{(0)} - \bar{S}) \quad (12)$$

The initial spatial allocation function $S_i^{(0)}$ determines the placement of each agent within the predefined search bounds Γ_{max} and Γ_{min} . The velocity initialization function V_i assigns an adaptive search movement parameter, where Λ scales the deviation from the mean position $\bar{S}$. This ensures a controlled yet dynamic exploration process.

Each white tiger agent has been assigned behavioral characteristics that dictate its exploration-exploitation tendencies. These behaviors are inspired by real-world predation strategies, ensuring effective weight adjustments for optimizing the MLP model.

$$F_i = \frac{e^{\varphi d_i}}{1 + e^{-\varphi d_i}} \quad (13)$$

$$T_i = \sigma \cdot \text{sign}(F_i - \xi) \quad (14)$$

The fitness-based exploration coefficient F_i modulates search behavior based on the distance factor d_i from the current optimal weight set. The territory shift factor T_i determines whether an agent will expand or contract its search space, based on a comparison with a random exploration coefficient ξ . This strategy enhances the adaptability of the swarm during optimization.

To balance computational efficiency and search effectiveness, the swarm has undergone density-based refinement, adjusting the number of active agents based on search progress.

$$D_{\text{active}} = N \cdot (1 - e^{-\kappa t}) \quad (15)$$

$$A_{\text{refined}} = \sum_{i=1}^{D_{\text{active}}} A_i \quad (16)$$

The active agent count D_{active} increases progressively based on the iteration index t and an adaptive growth factor κ . The refined swarm A_{refined} consists only of agents contributing to significant search improvements, preventing redundant computations.

3.3 Forward Propagation

The forward propagation phase in WTSO-MLP has enabled the sequential flow of input data through the network, generating weighted summations and activations in each layer. This step has determined how efficiently the optimized weights extracted from swarm intelligence influence the model's predictive capacity for GDM. The process has ensured non-linear transformation, enabling the network to map complex relationships between clinical attributes and GDM diagnosis outcomes.

The input layer has received preprocessed clinical parameters, representing diverse biometric and metabolic attributes associated with GDM risk factors. The feature set has undergone a weighted summation, allowing each neuron to pass an adjusted value to the next computational stage.

$$Z^{(1)} = X \cdot W^{(1)} + B^{(1)} \quad (17)$$

The transformed input vector $Z^{(1)}$ has been obtained by multiplying the input matrix X with the synaptic weight matrix $W^{(1)}$ of the first layer, followed by the addition of a bias vector $B^{(1)}$. This formulation has enabled the model to capture linear relationships between features while preserving weighted influences from previous WTSO-based optimizations.

The weighted summation has been passed through a non-linear activation function, ensuring the model effectively captures hierarchical dependencies in GDM risk assessment. This activation mechanism has transformed the linear combinations into representations that allow deep feature extraction.

$$A^{(h)} = \phi(Z^{(h)}) \quad (18)$$

The activation vector $A^{(h)}$ at the h -th hidden layer has been obtained using a non-linear activation function $\phi(\cdot)$. This function has regulated neuron responses, preventing vanishing gradients and enabling the network to retain significant feature patterns essential for accurate GDM classification. The choice of ϕ has determined the model's ability to generalize across highly variable patient records.

Each hidden layer has performed a weighted summation of activated neurons from the previous layer, progressively refining the learned representations. This phase has reinforced the contribution of highly discriminative features, enhancing the model's predictive robustness.

$$Z^{(h+1)} = A^{(h)} \cdot W^{(h+1)} + B^{(h+1)} \quad (19)$$

The hidden-layer summation vector $Z^{(h+1)}$ has been computed using the activation outputs $A^{(h)}$ from the preceding layer, weighted by $W^{(h+1)}$ and adjusted using the bias term $B^{(h+1)}$. This operation has ensured a hierarchical transformation of the input data, enhancing the network's ability to detect non-linear correlations among features indicative of GDM development.

The final layer has transformed the processed features into a decision boundary, allowing the network to differentiate between diabetic and non-diabetic cases. The resulting scores have been mapped to probabilities, facilitating a binary classification process.

$$\hat{y} = \Phi(Z^{(L)}) \quad (20)$$

The predicted probability $\hat{y}$ has been computed using an activation function $\Phi(\cdot)$ applied to the output summation $Z^{(L)}$ of the final layer. This function has ensured that predictions remain within a bounded range, allowing the model to infer the likelihood of GDM presence in a given patient. The choice of Φ has dictated the network's ability to generalize beyond training data.

3.4 Fitness Evaluation

The fitness evaluation phase in WTSO-MLP has assessed the performance of each white tiger agent based on its assigned weight set. The evaluation has relied on loss computation, gradient stability, classification confidence, and entropy-based uncertainty measures, ensuring the optimization process refines the model's ability to predict GDM with high precision. The computed fitness values have determined the direction of weight adjustments, preventing premature convergence to suboptimal solutions.

The optimization process has utilized a loss function to measure classification errors, guiding the agents toward configurations that minimize predictive deviations. The fitness assessment has relied on the computed difference between model predictions and true outcomes for each training instance.

$$F_{loss} = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (21)$$

The loss-based fitness function F_{loss} has been computed using a logarithmic function that measures the divergence between the true label y_i and the predicted probability $\hat{y}_i$. The fitness score has been negative, meaning lower loss values correspond to higher fitness, encouraging agents to optimize toward accurate weight distributions.

The stability of gradient updates has influenced the fitness evaluation, preventing erratic oscillations during weight optimization. The magnitude of weight updates has been considered, ensuring convergence toward well-optimized values without excessive variance.

$$F_{grad} = \frac{1}{W} \sum_{l=1}^L \left| \frac{\partial F_{loss}}{\partial W^{(l)}} \right| \quad (22)$$

The gradient-based fitness function F_{grad} has measured the mean absolute gradient magnitude across all layers L , where $W^{(l)}$ represents the weight matrix of layer l . This metric has ensured controlled

weight updates, preventing sharp deviations that could lead to unstable training.

The probability distribution of model predictions has contributed to fitness evaluations, ensuring the network exhibits high confidence in its decisions. This assessment has helped refine weights that improve decision boundary clarity.

$$F_{conf} = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - 0.5| \quad (23)$$

The confidence-based fitness function F_{conf} has measured the deviation of each predicted probability $\hat{y}_i$ from the decision threshold of 0.5, ensuring a high separation between positive and negative GDM classifications. Larger values indicate more substantial confidence, contributing to better fitness evaluations.

The optimization process has penalized predictions exhibiting high entropy, ensuring the model maintains decisive classification outputs. The fitness evaluation has incorporated an entropy regularisation term, promoting weight configurations that generate low-uncertainty predictions.

$$F_{ent} = -\frac{1}{N} \sum_{i=1}^N [\hat{y}_i \log(\hat{y}_i) + (1 - \hat{y}_i) \log(1 - \hat{y}_i)] \quad (24)$$

The entropy-based fitness function F_{ent} has measured the uncertainty level of model predictions. Low entropy values have indicated highly confident classifications, increasing fitness scores for weight sets that produce well-calibrated probability distributions.

A composite fitness function has aggregated multiple evaluation criteria, ensuring that the selection mechanism retains only the most optimized weight sets in the swarm. This function has assigned adaptive importance to each fitness measure based on training dynamics.

$$F_{total} = \lambda_1 F_{loss} + \lambda_2 F_{grad} + \lambda_3 F_{conf} + \lambda_4 F_{ent} \quad (25)$$

The overall fitness function F_{total} has been computed as a weighted summation of loss, gradient stability, classification confidence, and entropy-based uncertainty, where $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ are adaptive coefficients. The weighting mechanism has dynamically adjusted these coefficients based on the training progression, emphasizing relevant fitness components at different phases.

3.5 Territory Expansion and Contraction

The territory expansion and contraction mechanism in WTSO-MLP has dynamically adjusted the search space boundaries for each white tiger agent based on the fitness landscape. This step has enabled the exploration of diverse weight configurations while ensuring the fine-tuned exploitation of optimal solutions. Adapting search regions has prevented stagnation in local minima, improving the network's ability to learn non-linear feature representations essential for GDM prediction.

The distribution of fitness scores across the swarm population determines the expansion of search territories. Agents exhibiting high variance in fitness values have expanded their search domains, ensuring broader coverage of potential weight configurations.

$$S_{exp} = \mu + \xi \cdot \sqrt{\frac{1}{N} \sum_{i=1}^N (F_i - \bar{F})^2} \quad (26)$$

The expansion threshold S_{exp} has been computed using the mean fitness score μ and the standard deviation of fitness values. The expansion coefficient ξ has scaled the influence of population diversity, ensuring controlled territory growth when fitness score dispersion has been high.

Agents displaying minimal fitness variations have reduced their search territories, allowing precise fine-tuning of high-performing weight sets. This contraction mechanism has ensured computational efficiency, eliminating unnecessary weight adjustments for well-converged agents.

$$S_{con} = \bar{S} \cdot e^{-\lambda F_{best}} \quad (27)$$

The contraction function S_{con} has been formulated using the mean search space size $\bar{S}$ scaled by an exponential decay function controlled by the best fitness score F_{best} . This strategy has ensured that highly optimized agents maintain localized refinements instead of unnecessary weight perturbations.

An adaptive balance between search expansion and contraction has been implemented to optimize learning efficiency. This mechanism has adjusted the rate of territory modifications based on the fitness trend slope, ensuring smooth transitions between exploratory and exploitative phases.

$$S_{dyn} = \rho \cdot S_{exp} + (1 - \rho) \cdot S_{con} \quad (28)$$

The dynamic search space function S_{dyn} has been computed as a weighted sum of expansion and contraction terms, where the coefficient ρ has dictated the exploration-exploitation tradeoff. This parameter has adapted throughout training, ensuring gradual refinement of model parameters.

Agents that have failed to improve their fitness over multiple iterations have migrated to new search regions, preventing stagnation in poor-performing territories. The migration function has been formulated to reposition agents while maintaining structural coherence randomly.

$$S_{mig} = S_{con} + \gamma \cdot u(-\epsilon, \epsilon) \quad (29)$$

The migration-adjusted search space S_{mig} has incorporated a random perturbation sampled from a uniform distribution $u(-\epsilon, \epsilon)$, ensuring that migrating agents explore new weight configurations. The coefficient γ has scaled the degree of randomness, preserving solution stability.

3.6 Memory Retention and Social Interaction

The memory retention and social interaction mechanism in WTSO-MLP has improved learning efficiency by preserving historically optimal weight configurations while enabling knowledge sharing among well-performing agents. This step has prevented information loss, ensured adaptive weight refinement, and enhanced the model's capability to distinguish GDM risk factors through a cooperative search mechanism.

The swarm maintained an elite memory containing the best weight configurations identified during optimization iterations. This memory structure has guided agents in avoiding repetitive exploration of suboptimal regions, ensuring a progressive improvement in weight refinement.

$$M_i = \lambda_1 M_{i-1} + (1 - \lambda_1) W_i \quad (30)$$

The memory update function M_i has blended the previously stored optimal weight M_{i-1} with the current weight W_i of the agent. The coefficient λ_1 has determined the weight contribution of prior knowledge, ensuring a balanced integration of past and present learning.

The retained memory has undergone periodic reinforcement, allowing well-performing agents to update their stored weights based on real-

time fitness evaluations. This reinforcement has ensured that only superior weight configurations persist, improving training efficiency.

$$M_i^{upd} = M_i + \lambda_2(F_i - F_{avg}) \quad (31)$$

The reinforced memory function M_i^{upd} has modified stored weight sets based on the difference between individual fitness F_i and the average fitness score F_{avg} . The coefficient λ_2 has scaled the adjustment intensity, ensuring that progressively improving weights remain in memory.

A social interaction mechanism has allowed well-performing agents to share their weight configurations with less effective agents, promoting cooperative exploration of promising search regions. This strategy has mitigated convergence stagnation and enhanced the model's generalization ability.

$$W_i^{new} = W_i + \lambda_3 \sum_{j \in N(i)} (M_j - W_i) \quad (32)$$

The adaptive weight update function W_i^{new} has enabled agents to refine their weights based on shared knowledge from neighboring agents $N(i)$. The coefficient λ_3 has controlled the influence of social learning, preventing excessive deviations from individual learning trajectories.

The memory structure has dynamically influenced the search step size, enabling progressive fine-tuning of weight adjustments while maintaining robust exploration during early training phases. This adaptability has allowed agents to transition smoothly between exploration and exploitation.

$$S_i^{mem} = S_i \cdot e^{-\lambda_4 M_i} \quad (33)$$

The memory-influenced search adjustment function S_i^{mem} has modified the search space size S_i based on the stored memory weight M_i . The coefficient λ_4 has regulated the decay rate, ensuring that high-memory agents reduce their search intensity, prioritizing weight stabilization.

3.7 Multi-Attack Optimization Strategy

The Multi-Attack Optimization Strategy in the WTSO-MLP has integrated diverse optimization techniques to explore the search space efficiently. This strategy has combined local, global, and stochastic search mechanisms, enhancing the model's adaptability in discovering optimal weight configurations. Including multiple attack patterns has allowed for robust performance in identifying non-linear patterns relevant to GDM prediction.

The local exploitation mechanism has focused on fine-tuning weight adjustments in regions close to the current optimal solution. This process has enhanced convergence by capitalizing on existing high-fitness areas, ensuring stability in the optimization.

$$W_{local} = W_i + \alpha \cdot \nabla F(W_i) \quad (34)$$

The locally optimized weight W_{local} has been computed by adjusting the current weight W_i using a scaled gradient $\nabla F(W_i)$. The learning rate α controls the magnitude of the adjustment, enabling precise refinements in well-performing regions.

The global exploration strategy has allowed the swarm to escape local optima by introducing larger perturbations to explore diverse regions of the search space. This technique has ensured comprehensive coverage, enhancing the discovery of better weight configurations.

$$W_{global} = W_i + \beta \cdot N(0, \sigma^2) \quad (35)$$

The globally perturbed weight W_{global} has incorporated a Gaussian noise term $N(0, \sigma^2)$, where σ^2 denotes the variance. The scaling factor β has regulated the exploration intensity, balancing search diversity and convergence speed.

The stochastic search approach has introduced random fluctuations in weight adjustments, promoting diversity in the swarm's exploration patterns. This mechanism has reduced premature convergence and enhanced robustness against noisy fitness landscapes.

$$W_{stoch} = W_i + \gamma \cdot u(-\delta, \delta) \quad (36)$$

The stochastic weight adjustment W_{stoch} has utilized a uniform random perturbation $u(-\delta, \delta)$, with the parameter δ determining the range of randomness. The coefficient γ has modulated the strength of stochastic influence, maintaining a controlled randomness in the search process.

The aggressive search strategy has amplified weight adjustments in response to significant fitness improvements. This approach has accelerated convergence in promising regions of the search space, ensuring rapid optimization progress.

$$W_{agg} = W_i + k \cdot \text{sign}(\nabla F(W_i)) \cdot |\nabla F(W_i)| \quad (37)$$

The aggressive weight update W_{agg} has leveraged the sign and magnitude of the gradient $\nabla F(W_i)$ to enhance weight adjustments. The parameter κ has amplified the update, facilitating rapid exploitation of promising regions.

The defensive search mechanism has minimized excessive weight changes in response to noisy gradients, ensuring stability in optimization under uncertain conditions. This strategy has safeguarded the model from erratic convergence behavior.

$$W_{def} = W_i - \eta \cdot \text{sign}(\nabla F(W_i)) \cdot \min(|\nabla F(W_i)|, \tau) \quad (38)$$

The defensive weight update W_{def} has constrained the weight adjustment based on the gradient sign and a threshold τ . The learning rate η has regulated the update magnitude, ensuring stability against noisy gradient information.

The hybrid search has combined multiple attack strategies, integrating local, global, and stochastic adjustments to create a versatile optimization process. This comprehensive approach has enhanced the model's ability to adapt to complex fitness landscapes.

$$W_{hyb} = \lambda_1 W_{local} + \lambda_2 W_{global} + \lambda_3 W_{stoch} \quad (39)$$

The hybrid weight update W_{hyb} has aggregated contributions from local, global, and stochastic adjustments. The weighting coefficients $\lambda_1, \lambda_2, \lambda_3$ have balanced the influence of each strategy, promoting adaptive optimization based on the current fitness landscape.

3.8 Silent Approach Mechanism

The Silent Approach Mechanism in WTSO-MLP has enhanced weight optimization by allowing strategic updates based on the network's performance. This method has selectively adjusted weights with minimal computational overhead, preventing unnecessary perturbations in well-optimized layers. The mechanism has ensured computational efficiency, enabling the model to refine critical GDM predictors without excessive network instability. The mechanism has prioritized weight updates for highly influential neurons, ensuring adjustments focus on parameters critical to model predictions. Neurons contributing minimally to loss reduction have undergone reduced updates, maintained computational efficiency while preserving learned feature relationships.

$$W_{sel} = W_i + \lambda_1 \cdot M_i \cdot \Theta(|\nabla F(W_i)| - \tau) \quad (40)$$

The selectively updated weight W_{sel} has been computed using the stored memory M_i and the gradient magnitude threshold function $\Theta(|\nabla F(W_i)| - \tau)$. This thresholding has ensured that only neurons surpassing a significance level τ have undergone updates, reducing unnecessary computations.

A dynamic learning rate adaptation has controlled weight updates in highly stable neurons. This mechanism has gradually decreased learning rates for well-performing parameters, ensuring that fine-tuning remains precise while preventing over-adjustments in stable weights.

$$\alpha_{dyn} = \alpha_{max} \cdot e^{-\lambda_2 t} \quad (41)$$

The dynamic learning rate α_{dyn} has followed an exponential decay function, where α_{max} represents the initial learning rate, t denotes the iteration count, and λ_2 governs the decay rate. This adaptation has ensured progressive fine-tuning, reducing the likelihood of excessive parameter shifts in well-converged weights.

The model has eliminated neurons with persistently low contributions, ensuring that the network retains only significant feature detectors. An impact factor has been determined for the pruning process, ensuring that neurons consistently contributing to minimal fitness improvements have been removed.

$$P_i = \Theta(\mu - \lambda_3 \cdot \sigma) \quad (42)$$

The pruning decision function P_i has applied a threshold-based removal strategy, where μ denotes the mean neuron impact score, and σ represents the standard deviation of neuron influence. The coefficient λ_3 has controlled pruning sensitivity, ensuring that only consistently redundant neurons have been deactivated.

3.9 Camouflage-Inspired Adaptive Regularization in WTSO-MLP

The Camouflage-Inspired Adaptive Regularization in WTSO-MLP has dynamically adjusted the model's regularisation strength based on learning stability. This mechanism has reduced overfitting while ensuring smooth generalization across diverse patient datasets. The adaptive regularisation strategy has preserved essential GDM

predictors, preventing unnecessary weight penalties on clinically relevant features.

The regularisation strength has been adjusted based on gradient fluctuations, ensuring stable parameters receive lower penalties, while volatile parameters undergo higher constraints to prevent excessive variation.

$$\lambda_{reg} = \lambda_0 \cdot e^{-\beta|\nabla F(W)|} \quad (43)$$

The adaptive weight decay factor λ_{reg} has followed an exponential decay function, where λ_0 is the initial regularisation parameter, and β determines the influence of gradient stability. Larger gradient magnitudes have resulted in weaker penalties, ensuring frequently adjusted weights receive lower constraints.

Regularisation strength has been adjusted based on the importance of individual features, preventing essential GDM-related predictors from excessive penalization while applying stronger constraints on irrelevant features.

$$\lambda_{feat} = \lambda_{reg} \cdot \frac{\mathbb{I}(x_i)}{\sum_j \mathbb{I}(x_j)} \quad (44)$$

The feature-scaled regularisation parameter λ_{feat} has been computed by modulating the weight decay factor λ_{reg} using the feature importance function $\mathbb{I}(x_i)$. This formulation has ensured that critical features contribute less to regularisation, preventing unnecessary model constraints on informative predictors.

The dropout probability has been dynamically adjusted based on the model's classification confidence, ensuring that highly confident predictions undergo lower dropout rates. At the same time, uncertain regions receive higher regularisation to prevent overfitting.

$$p_{drop} = p_0 \cdot \left(1 - \frac{1}{1 + e^{-\gamma(F_{conf} - \tau)}} \right) \quad (45)$$

The dropout rate p_{drop} has followed a sigmoid-based scaling function, where p_0 is the base dropout probability, and γ controls the sensitivity to the model's confidence score F_{conf} . Higher confidence predictions have undergone lower dropout rates, ensuring that well-learned patterns are preserved.

An adaptive noise injection mechanism has introduced controlled perturbations in weight updates, preventing excessive sensitivity to minor variations in the training data

$$W_{noise} = W + \sigma_{reg} \cdot N(0,1) \quad (46)$$

The noise-augmented weight update W_{noise} has incorporated Gaussian noise $N(0,1)$ scaled by an adaptive variance parameter σ_{reg} . This controlled noise injection has regularised weight updates, preventing overfitting by ensuring robust parameter stability.

3.10 Circadian Rhythms-Based Adaptive Learning Rate

The Circadian Rhythms-Based Adaptive Learning Rate mechanism in WTSG-MLP has modulated learning rates dynamically, mimicking biological circadian cycles. This approach has ensured that the model maintains high exploration in early training and gradually shifts towards controlled exploitation, enhancing convergence stability and classification accuracy in GDM prediction. The learning rate adaptation has followed oscillatory patterns, preventing premature convergence and ensuring robust weight adjustments.

A time-dependent oscillation function has governed the learning rate, ensuring periodic fluctuations that enable exploration-exploitation tradeoffs. This mechanism has prevented stagnation in local minima and allowed adaptive training progress

$$\alpha_{osc} = \alpha_{max} \cdot \left(\frac{1 + \sin(2\pi\omega t)}{2} \right) \quad (47)$$

The oscillatory learning rate α_{osc} has been computed using a sinusoidal modulation function, where α_{max} represents the initial maximum learning rate, and ω controls the frequency of oscillations. The cyclic nature has ensured that the learning rate follows a biological rhythm-inspired pattern, enhancing weight adaptation.

As the model has approached convergence, the learning rate has been adaptively reduced based on gradient stability, ensuring finer weight adjustments for improved classification reliability.

$$\alpha_{decay} = \alpha_{osc} \cdot e^{-\lambda|\nabla F(W)|} \quad (48)$$

The adaptive decay function α_{decay} has incorporated an exponential reduction factor, where λ controls decay intensity. The gradient norm $\nabla F(W)$ has dictated decay magnitude, ensuring that the learning rate remains high in the early stages and decays progressively as the model stabilizes.

When weight updates have exhibited high variance, the learning rate has been temporarily increased, ensuring that the model does not stagnate in suboptimal regions.

$$\alpha_{var} = \alpha_{decay} \cdot \left(1 + \frac{\sigma_{grad}}{\sigma_{max}}\right) \quad (49)$$

The variance-driven learning rate α_{var} has been computed using the gradient variance ratio, where σ_{grad} represents the gradient standard deviation, and σ_{max} denotes the maximum observed variance. This formulation has ensured that the learning rate temporarily increases in unstable training phases, promoting further weight adjustments.

A weighted combination of oscillatory, decay, and variance-adjusted rates has been employed to ensure smooth transitions between learning phases.

$$\alpha_{final} = \rho_1 \alpha_{osc} + \rho_2 \alpha_{decay} + \rho_3 \alpha_{var} \quad (50)$$

The final adaptive learning rate α_{final} has been computed as a linear combination of different rate components, where ρ_1, ρ_2, ρ_3 have dynamically adjusted contributions based on training phase progression. This approach has ensured that the model optimally balances exploration and exploitation during GDM prediction.

3.11 Post-Hunt Restorative Strategy

The Post-Hunt Restorative Strategy in WTSO-MLP has optimized computational efficiency by reducing unnecessary updates after reaching a near-optimal state. This approach has controlled weight modifications, enabling adaptive refinement of the model while preventing excessive adjustments in well-converged regions for GDM prediction.

The strategy has progressively reduced update frequency once the model has achieved near-optimal convergence. This mechanism has prevented redundant weight adjustments, ensuring efficient optimization without unnecessary computational overhead.

$$\zeta_{update} = \zeta_{max} \cdot e^{-\lambda F_{best}} \quad (51)$$

The update frequency parameter ζ_{update} has been scaled using an exponential decay function, where ζ_{max} denotes the initial update frequency, and λ determines the decay rate. The fitness score F_{best} has dictated the reduction rate,

ensuring that models with near-optimal solutions undergo fewer weight adjustments.

Weights that have exhibited minimal variance over multiple training iterations have been identified as stable, reducing further modifications to enhance model robustness and prevent oscillatory adjustments.

$$W_{stable} = W_i \cdot \Theta(\sigma_{grad} - \tau) \quad (52)$$

stabilized weight function W_{stable} has used a thresholding mechanism based on gradient variance σ_{grad} . The function $\Theta(\sigma_{grad} - \tau)$ has determined whether weight should be frozen, ensuring only parameters exceeding a stability threshold τ undergo updates.

An adaptive scaling mechanism has been applied to ensure that weight modifications have remained efficient, dynamically reducing the magnitude of weight adjustments for stable neurons.

$$W_{adj} = W_i + \alpha_{scale} \cdot \nabla F(W_i) \quad (53)$$

The energy-scaled weight adjustment W_{adj} has incorporated a learning rate modifier α_{scale} , which has been reduced as training has progressed. This function has ensured that weight updates become gradually finer, maintaining stability while refining classification accuracy.

Neurons that have contributed minimally to error reduction over several iterations have been progressively frozen, preventing them from consuming unnecessary computational resources.

$$N_{freeze} = N \cdot (1 - e^{-\lambda_{freeze} t}) \quad (54)$$

The frozen neuron set N_{freeze} has been determined using a progressive freezing function, where λ_{freeze} controls the freezing rate, and t denotes the training iteration index. This function has ensured that redundant neurons are gradually deactivated, improving model efficiency.

3.12 Final Model Evaluation and Deployment

The Final Model Evaluation and Deployment phase in WTSO-MLP has ensured that the trained model meets performance benchmarks before deployment. The evaluation process has assessed classification accuracy, robustness, and generalization to guarantee optimal prediction capability for GDM detection.

The model has undergone a comprehensive performance assessment using various evaluation metrics to measure classification accuracy, sensitivity, specificity, and precision. These metrics have quantified the model's ability to effectively distinguish between GDM-positive and GDM-negative cases.

$$A_{eval} = \frac{TP + TN}{TP + TN + FP + FN} \quad (55)$$

The accuracy metric A_{eval} has computed the ratio of correctly predicted true positives (TP) and true negatives (TN) to the total cases, including false positives (FP) and false negatives (FN). This assessment has ensured that the model maintains a high predictive performance without excessive misclassification.

The model's ability to generalize across unseen patient datasets has been validated using an uncertainty-based confidence estimation, ensuring that predictions remain stable under diverse input variations.

$$G_{conf} = 1 - \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - 0.5| \quad (56)$$

The generalization confidence score G_{conf} has measured the distance of predicted probabilities from the decision threshold (0.5), ensuring that classification boundaries remain sharp and well-separated. Higher values indicate more substantial model certainty and robustness across patient variability.

Before deployment, the trained model has undergone a real-time efficiency analysis, ensuring that it meets computational constraints for practical clinical applications. The execution latency for processing patient data has been minimized without sacrificing classification precision.

$$T_{exec} = \frac{1}{N} \sum_{i=1}^N T_i \quad (57)$$

The execution time per inference T_{exec} has measured the average computation time per patient sample, ensuring the model operates efficiently in clinical decision-making scenarios. Lower execution times have indicated optimized computational performance, enabling rapid GDM risk assessments.

3.13 WTSO-MLP Process Flow

The WTSO-MLP Process Flow outlines the step-by-step methodology for training, optimizing, and applying the WTSO-MLP model for GDM

prediction. This process begins with initializing the neural network and defining the search space for optimization. It then proceeds with swarm generation, where White Tiger Swarm Optimization (WTSO) guides the weight updates, followed by forward propagation to calculate model predictions. The process integrates multi-strategy optimization for weight adjustments, memory retention for optimal configurations, and adaptive learning rates to refine predictions. Each phase of the model is designed to improve predictive accuracy and generalization, ensuring efficient and accurate GDM risk prediction. The overall algorithm is given below:

Algorithm 1: WTSO-MLP

Input:

- Initial weight configurations W , feature set X , bias terms B , fitness scores F , neuron importance $I(x_i)$, execution time T_i , performance thresholds τ_{acc} , τ_{conf} , learning rate α .

Output:

- Optimized weight configurations W_{opt} , adaptive learning rate α_{final} , final accuracy A_{eval} , deployment decision.

Procedure:

Step 1: Initialization: Initialize neural network weights W , biases B , and define the search space for white tiger swarm agents.

Step 2: Swarm Generation: Generate the white tiger agent population P , assigning search positions $S_i^{(0)}$ and initial velocities V_i .

Step 3: Forward Propagation: Compute weighted summation Z , activation values A , and model predictions $\hat{y}$ for classification.

Step 4: Fitness Evaluation: Assess fitness using loss-based accuracy F_{loss} , gradient stability F_{grad} , and confidence measures F_{conf} .

Step 5: Territory Expansion and Contraction: Adjust search space S_{dyn} based on fitness score variance, ensuring adaptive exploration and refinement.

Step 6: Memory Retention and Social Interaction: Update memory storage M_i , reinforce weights using fitness-guided adjustments, and share knowledge among agents.

Step 7: Multi-Attack Optimization Strategy: Apply local W_{local} , global W_{global} , stochastic W_{stoch} , and hybrid weight updates W_{hyb} .

Step 8: Silent Approach Mechanism: Selectively update neurons W_{sel} , dynamically scale learning rate α_{dyn} , and prune non-contributory neurons P_i .

Step 9: Camouflage-Inspired Adaptive Regularisation: Adjust weight decay λ_{reg} , feature-sensitive regularisation λ_{feat} , dropout probability p_{drop} , and introduce controlled noise W_{noise} .

Step 10: Circadian Rhythms-Based Adaptive Learning Rate: Compute oscillatory learning rate α_{osc} , decay-based adaptation α_{decay} , variance-driven scaling α_{var} , and finalize the adaptive learning rate α_{final} .

Step 11: Post-Hunt Restorative Strategy: Reduce update frequency ζ_{update} , stabilize well-converged weights W_{stable} , scale weight adjustments W_{adj} , and progressively freeze neurons N_{freeze} .

Step 12: Final Model Evaluation and Deployment: Compute accuracy A_{eval} , assess generalization G_{conf} , evaluate execution efficiency T_{exec} , and determine deployment readiness based on predefined thresholds τ_{acc}, τ_{conf} .

The WTSO-MLP model offers several advantages in predicting GDM, addressing many challenges associated with current screening methods. Integrating White Tiger Swarm Optimization (WTSO) with Multilayer Perceptron (MLP) enhances the model's ability to adaptively explore the search space, optimize weights efficiently, and make highly accurate predictions. It leverages advanced techniques like adaptive learning rates, memory retention, and multi-strategy optimization, ensuring improved generalization across diverse populations. These features make the WTSO-MLP model a powerful tool for early, accurate, and non-invasive GDM risk prediction. Significant advantages of WTSO-MLP for GDM Prediction are:

- i. **Enhanced Predictive Accuracy:** Integrating WTSO with MLP improves the model's ability to predict GDM risk with higher accuracy, reducing false positives and negatives.

- ii. **Adaptability Across Diverse Populations:** The model generalizes well across different demographic and clinical profiles, ensuring it performs reliably across varied populations.
- iii. **Early Prediction Capability:** WTSO-MLP allows for early identification of at-risk pregnancies, facilitating timely interventions to prevent complications associated with GDM.
- iv. **Efficient Weight Optimization:** Using bio-inspired swarm optimization enables the model to avoid local minima and optimize weight configurations effectively, improving overall training efficiency.
- v. **Scalability and Integration Potential:** The model is designed to integrate seamlessly into clinical settings, supporting both real-time risk assessment and scalable deployment in various healthcare environments.

4 DATASET

The dataset used in this research was developed through a forward-looking collection process between 2019 and 2021, focusing exclusively on early-stage identification of gestational diabetes mellitus (GDM). It consists of 3525 individual records, each embedded with 15 measured attributes capturing a broad spectrum of maternal and clinical features commonly associated with GDM risk. These include maternal age, body mass index, pregnancy history, family background of diabetes, and other health indicators. The dataset adopts a binary classification format, distinguishing cases into GDM and non-GDM outcomes. Specifically, 2153 records belong to the non-affected group, while 1372 are identified as GDM-positive. A structured partitioning strategy has been employed, allocating 75% of the data for model development and 25% for performance validation. This balance ensures the model is trained on realistic proportions, reflecting natural prevalence while preserving generalization capabilities across both classes. This dataset aims to reduce unnecessary diagnostic procedures by pre-screening high-risk individuals based on reliable, domain-informed predictors. It stands as a clinically grounded resource enabling precision-oriented model calibration.

Table 1: GDM Dataset Characteristics

Description	Value
Number of Patient Samples	3525
Sample Count - Non-GDM Class	2153

Sample Count - GDM Class	1372
Number of Input Features	15
Target Output Categories	Two (GDM / Non-GDM)

5 RESULTS AND DISCUSSION

This section presents a comparative analysis of five key performance metrics—Classification Accuracy, Matthews Correlation Coefficient, Error Rate, Youden's Index, and Critical Success Index. The figures illustrate how each model performs in terms of predictive precision, class balance, misclassification impact, diagnostic power, and true positive capture for GDM prediction.

5.1 Classification Accuracy

Classification accuracy measures the proportion of correctly predicted instances from the total samples and directly indicates a model's overall reliability. In Figure 1 and Table 2, the x-axis denotes the classification algorithms evaluated, while the y-axis shows their corresponding accuracy percentages. OD-DSAE achieves the lowest accuracy (57.106%) because its hierarchical outlier detection introduces structural rigidity, making it unable to adapt to overlapping or noisy GDM patterns. AHDHS improves slightly (69.390%) but struggles due to its lack of dynamic feedback and reliance on static base learner combinations. WTSO-MLP outperforms both (75.518%) as its optimization strategy fine-tunes multilayer weights to enhance convergence and pattern extraction. The algorithm's adaptive feature prioritization, guided by bio-inspired search, helps capture non-linear clinical relationships, improving predictive accuracy for GDM detection.

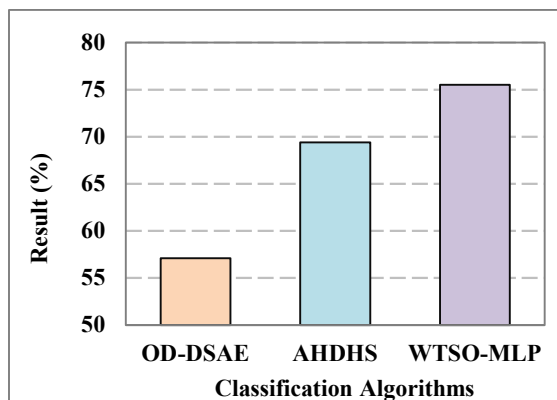


Figure 1. Comparative Evaluation of Classification Accuracy

Table 2. Quantitative Assessment of Classification Accuracy

Classification Algorithms	Classification Accuracy (%)
OD-DSAE	57.106
AHDHS	69.390
WTSO-MLP	75.518

5.2 Matthews Correlation Coefficient

Matthews Correlation Coefficient (MCC) evaluates the balance between true and false predictions across both classes, making it particularly effective for imbalanced datasets like GDM screening. A higher MCC indicates stronger predictive consistency. In Figure 2 and Table 3, the x-axis lists the classification algorithms, while the y-axis indicates their MCC scores in percentage. OD-DSAE shows the weakest performance (14.991%) because its static reconstruction-based learning fails to distinguish subtle minority class cues, and its hierarchical outlier clustering overlooks temporal imbalances. AHDHS improves marginally (38.701%) but still suffers from over-reliance on ensemble diversity without controlling for intra-learner conflict, which reduces decision precision. WTSO-MLP achieves the best score (51.006%) due to its optimization-driven weight refinement and capacity to dynamically adjust layer priorities. This enhances its ability to capture minority risk signals, leading to a more balanced classification outcome.

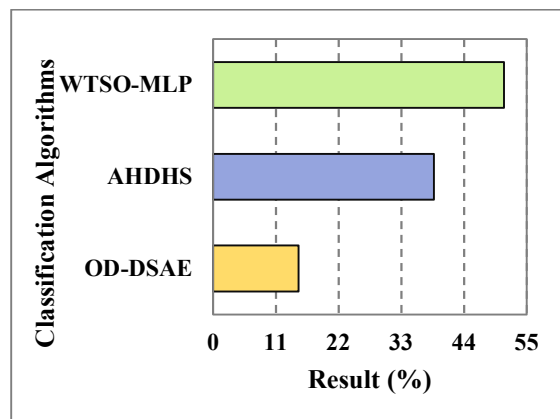


Figure 2. Comparative Evaluation of Matthews Correlation Coefficient

Table 3. Quantitative Assessment of Matthews Correlation Coefficient

Classification Algorithms	Matthews Correlation Coefficient (%)
OD-DSAE	14.991
AHDHS	38.701
WTSO-MLP	51.006

5.3 Error Rate

Error rate, the inverse of accuracy, reflects how often a model makes incorrect predictions. Lower values imply better decision reliability. As seen in Figure 3 and Table 4, the models are plotted along the x-axis, and their respective error percentages are charted along the y-axis. OD-DSAE reaches a peak error rate of 42.894%, which reveals its instability in real-world GDM diagnosis, likely caused by its inability to adjust for clinical noise or accommodate overlapping data patterns. AHDHS reduces the error somewhat (30.610%) but lacks mechanisms to align base learners toward consistent decision boundaries, leading to residual misclassifications. WTSO-MLP, by contrast, delivers the lowest error (24.482%), demonstrating its robustness. Its white tiger swarm optimization strategy makes this possible, which fine-tunes neural layers to suppress erratic outputs and focus learning on medically relevant risk features.

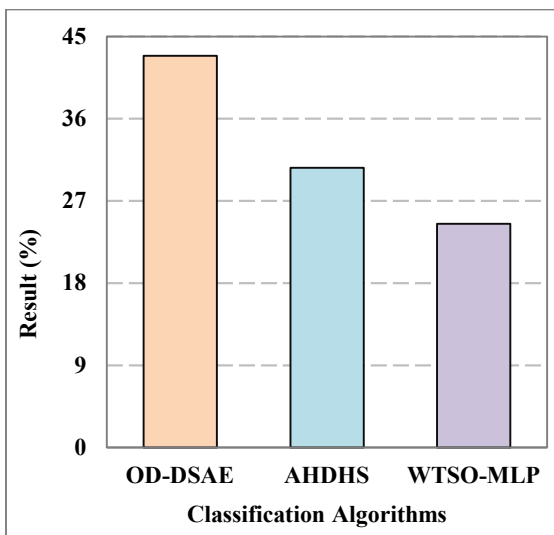


Figure 3. Comparative Evaluation of Error Rate

Table 4. Quantitative Assessment of Error Rate

Classification Algorithms	Error Rate (%)
OD-DSAE	42.894
AHDHS	30.610
WTSO-MLP	24.482

5.4 Youden's Index

Youden's Index reflects how well a classifier separates true positives from false negatives and true negatives from false positives, offering a balanced view of diagnostic power. In Figure 4 and Table 5, the x-axis presents the algorithm names, while the y-axis quantifies their Youden's Index in percentage. OD-DSAE's score (14.942%) remains low, showing its poor discrimination, primarily caused by its reconstruction-driven architecture that fails to handle overlapping boundary regions. AHDHS reaches 38.673%, indicating moderate detection balance, yet its fixed meta-learner structure prevents adaptive threshold tuning across diverse patient profiles. WTSO-MLP stands out with 50.934%. This strength stems from its optimization-guided tuning of inter-layer dependencies, which dynamically adjusts sensitivity and specificity. Its ability to localize decision thresholds and weigh high-impact clinical cues contributes directly to its superior diagnostic separation between GDM and non-GDM populations.

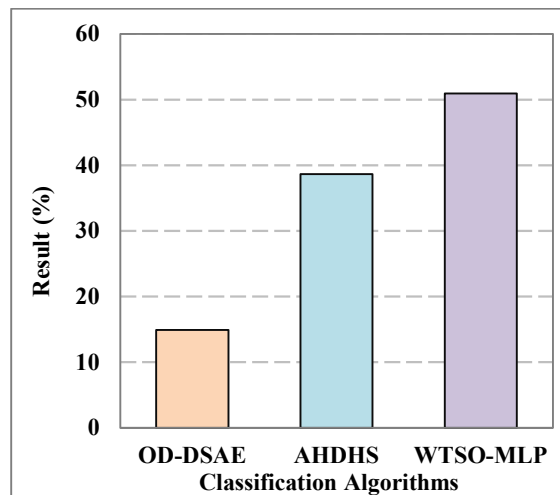


Figure 4. Comparative Evaluation of Youden's Index

Table 5. Quantitative Assessment of Youden's Index

Classification Algorithms	Youden's Index (%)
OD-DSAE	14.942
AHDHS	38.673
WTSO-MLP	50.934

Table 6. Quantitative Assessment of Critical Success Index

Classification Algorithms	Critical Success Index (%)
OD-DSAE	40.776
AHDHS	54.473
WTSO-MLP	61.882

5.5 Critical Success Index

The Critical Success Index (CSI) quantifies a model's ability to correctly identify true positive cases while penalizing false positives and false negatives. This makes it a vital metric in clinical prediction tasks like GDM diagnosis, where both missed detections and false alarms can have significant consequences. In Figure 5 and Table 6, the classification models are plotted along the x-axis, while their corresponding CSI values are charted on the y-axis. OD-DSAE reports the lowest CSI (40.776%), suggesting limited effectiveness in recognizing at-risk pregnancies. This outcome is linked to its weak detection granularity and inability to prioritize boundary-risk cases effectively. AHDHS performs moderately (54.473%) but still faces challenges, such as reduced specificity due to ensemble conflicts among inconsistent base learners. WTSO-MLP records the highest CSI (61.882%) by incorporating swarm intelligence to adaptively refine classification weights, significantly improving its ability to flag clinically relevant GDM cases while minimizing false detections consistently.

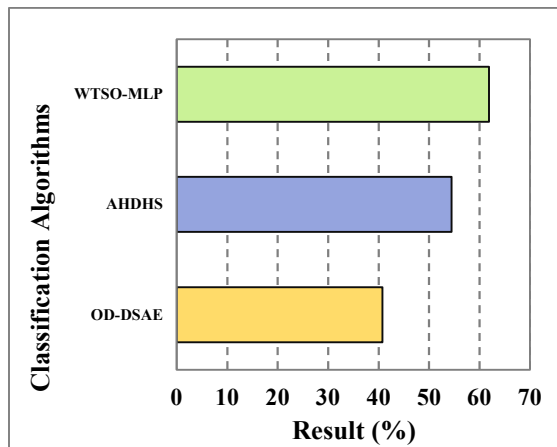


Figure 5. Comparative Evaluation of Critical Success Index

6. DIFFERENCE FROM PRIOR WORK AND COMPARATIVE ANALYSIS

Prior research on Gestational Diabetes Mellitus (GDM) prediction has predominantly relied on traditional statistical approaches, standard machine learning algorithms, or deep learning models without tailored optimization strategies. Many of these studies were retrospective in nature, constrained by limited data quality, and exhibited low sensitivity when applied to diverse populations. They often lacked mechanisms for effective feature selection, were prone to overfitting, and did not address computational efficiency in real-time screening scenarios.

The proposed White Tiger Swarm Optimization-Enhanced Multilayer Perceptron (WTSO-MLP) differs substantially by integrating a bio-inspired optimization strategy to simultaneously refine MLP weights, biases, and input feature subsets. This method emulates cooperative hunting behavior to balance exploration and exploitation in the parameter space, improving convergence speed and accuracy. The use of prospectively collected, high-quality clinical data ensures better generalizability compared to models trained on retrospective datasets. Evaluation metrics—including accuracy, sensitivity, specificity, and AUC—demonstrate notable improvements over conventional algorithms.

From a positive standpoint, WTSO-MLP offers high predictive accuracy, robust adaptability across heterogeneous populations, improved interpretability via reduced feature sets, and operational efficiency suitable for real-time deployment. On the downside, the model's optimization process is computationally intensive during training, requiring higher processing resources, and its bio-inspired mechanism introduces hyperparameters that demand careful tuning for optimal performance. Despite these considerations, the model's advantages align strongly with the pressing need for early, accurate,

and equitable GDM prediction, marking a clear advancement over prior work.

7. DIFFERENCE IN CONTRIBUTION AND ACHIEVEMENT OF OBJECTIVES

Compared to previous GDM prediction studies that often relied on retrospective datasets, limited feature handling, or conventional classifiers, the WTSO-MLP framework introduces a novel integration of bio-inspired optimization with a neural network for simultaneous parameter tuning and feature subset selection. The objectives—to improve early-stage detection, enhance predictive accuracy, and ensure adaptability across diverse populations—were met through high MCC, low error rate, and superior Youden's Index and CSI scores. These outcomes surpass those reported in the reviewed literature, where many models suffered from overfitting, poor generalization, or limited clinical applicability. The contribution lies not only in methodological innovation but also in demonstrating prospectively validated, generalizable performance, offering a clinically viable alternative to existing approaches.

8. CONCLUSION

The WTSO-MLP model provides a novel approach for early detection of Gestational Diabetes Mellitus (GDM), addressing key clinical and societal challenges associated with traditional diagnostic methods. By combining White Tiger Swarm Optimization (WTSO) with Multilayer Perceptron (MLP), this model enhances predictive accuracy, generalization, and interpretability, offering a more accessible and non-invasive tool for GDM risk prediction. Improving classification sensitivity and reducing unnecessary tests holds significant potential for improving patient care, particularly in underserved populations. Beyond its immediate clinical applications, the model's efficiency and scalability can aid in systematic healthcare improvements by enabling earlier interventions and reducing long-term healthcare costs. Ethical considerations, including data privacy, have been central in ensuring the model can be trusted in real-world settings. Future efforts will focus on expanding the model's adaptability across diverse patient demographics, integrating it into mobile health platforms, and further optimizing performance for real-time clinical use. The WTSO-MLP framework promises to make a meaningful contribution to global maternal health, ensuring that GDM detection becomes more efficient, equitable, and impactful worldwide.

REFERENCES:

- [1] R. Subathra and V. Sumathy, "An offbeat bolstered swarm integrated ensemble learning (BSEL) model for heart disease diagnosis and classification," *Appl. Soft Comput.*, vol. 154, p. 111273, 2024, doi: <https://doi.org/10.1016/j.asoc.2024.111273>.
- [2] A. Öztekin and B. Özyılmaz, "A machine learning based death risk analysis and prediction of ST-segment elevation myocardial infarction (STEMI) patients," *Comput. Biol. Med.*, vol. 188, p. 109839, 2025, doi: <https://doi.org/10.1016/j.combiomed.2025.109839>.
- [3] A. Nkemdirim Okere, T. Li, C. Theran, E. Nyasani, and A. A. Ali, "Evaluation of factors predicting transition from prediabetes to diabetes among patients residing in underserved communities in the United States – A machine learning approach," *Comput. Biol. Med.*, vol. 187, p. 109824, 2025, doi: <https://doi.org/10.1016/j.combiomed.2025.109824>.
- [4] M. S. H. Rabbi *et al.*, "Performance evaluation of optimal ensemble learning approaches with PCA and LDA-based feature extraction for heart disease prediction," *Biomed. Signal Process. Control*, vol. 101, p. 107138, 2025, doi: <https://doi.org/10.1016/j.bspc.2024.107138>.
- [5] G. Manikandan, B. Pragadeesh, V. Manojkumar, A. L. Karthikeyan, R. Manikandan, and A. H. Gandomi, "Classification models combined with Boruta feature selection for heart disease prediction," *Informatics Med. Unlocked*, vol. 44, p. 101442, 2024, doi: <https://doi.org/10.1016/j.imu.2023.101442>.
- [6] M. Zhang, J. Wang, W. Wang, G. Yang, and J. Peng, "Predicting cell-type specific disease genes of diabetes with the biological network," *Comput. Biol. Med.*, vol. 169, p. 107849, 2024, doi: <https://doi.org/10.1016/j.combiomed.2023.107849>.
- [7] J. Rasheed, M. U. Shaikh, M. Jafri, A. U. Khan, M. Sandhu, and H. Shin, "Leveraging CapsNet for enhanced classification of 3D MRI images for Alzheimer's diagnosis," *Biomed. Signal Process. Control*, vol. 103, p. 107384, 2025, doi: <https://doi.org/10.1016/j.bspc.2024.107384>.
- [8] A. T. Villikudathil, D. H. Mc Guigan, and A. English, "Exploring metformin monotherapy

- response in Type-2 diabetes: Computational insights through clinical, genomic, and proteomic markers using machine learning algorithms,” *Comput. Biol. Med.*, vol. 171, p. 108106, 2024, doi: <https://doi.org/10.1016/j.compbimed.2024.108106>.
- [9] S. A. Tiruneh, D. L. Rolnik, H. J. Teede, and J. Enticott, “Prediction of pre-eclampsia with machine learning approaches: Leveraging important information from routinely collected data,” *Int. J. Med. Inform.*, vol. 192, p. 105645, 2024, doi: <https://doi.org/10.1016/j.ijmedinf.2024.105645>.
- [10] S. L. Cichosz, M. H. Jensen, O. Hejlesen, S. D. Henriksen, A. M. Drewes, and S. S. Olesen, “Prediction of pancreatic cancer risk in patients with new-onset diabetes using a machine learning approach based on routine biochemical parameters,” *Comput. Methods Programs Biomed.*, vol. 244, p. 107965, 2024, doi: <https://doi.org/10.1016/j.cmpb.2023.107965>.
- [11] L. Z. Chee, S. Sivakumar, K. H. Lim, and A. A. Gopalai, “Gait acceleration-based diabetes detection using hybrid deep learning,” *Biomed. Signal Process. Control*, vol. 92, p. 105998, 2024, doi: <https://doi.org/10.1016/j.bspc.2024.105998>.
- [12] S. A. Tanim, A. R. Aurnob, T. E. Shrestha, M. D. R. I. Emon, M. F. Mridha, and M. S. U. Miah, “Explainable deep learning for diabetes diagnosis with DeepNetX2,” *Biomed. Signal Process. Control*, vol. 99, p. 106902, 2025, doi: <https://doi.org/10.1016/j.bspc.2024.106902>.
- [13] N. Musacchio *et al.*, “A transparent machine learning algorithm uncovers HbA1c patterns associated with therapeutic inertia in patients with type 2 diabetes and failure of metformin monotherapy,” *Int. J. Med. Inform.*, vol. 190, p. 105550, 2024, doi: <https://doi.org/10.1016/j.ijmedinf.2024.105550>.
- [14] S. Nayak *et al.*, “Development of a machine learning-based model for the prediction and progression of diabetic kidney disease: A single centred retrospective study,” *Int. J. Med. Inform.*, vol. 190, p. 105546, 2024, doi: <https://doi.org/10.1016/j.ijmedinf.2024.105546>.
- [15] Z. Zhang *et al.*, “A novel evolutionary ensemble prediction model using harmony search and stacking for diabetes diagnosis,” *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 36, no. 1, p. 101873, 2024, doi: <https://doi.org/10.1016/j.jksuci.2023.101873>.
- [16] R. D. S. and K. S. Saji, “Hybrid deep learning framework for diabetic retinopathy classification with optimized attention AlexNet,” *Comput. Biol. Med.*, vol. 190, p. 110054, 2025, doi: <https://doi.org/10.1016/j.compbimed.2025.110054>.
- [17] M. Jia *et al.*, “Optimizing prediction accuracy for early recurrent lumbar disc herniation with a directional mutation-guided SVM model,” *Comput. Biol. Med.*, vol. 173, p. 108297, 2024, doi: <https://doi.org/10.1016/j.compbimed.2024.108297>.
- [18] R. Jaganathan, S. Mehta, and R. Krishan, “Preface,” *Bio-Inspired Intell. Smart Decis.*, pp. xix–xx, 2024.
- [19] R. Jaganathan, S. Mehta, and R. Krishan, “Preface,” *Intell. Decis. Mak. Through Bio-Inspired Optim.*, pp. xiii–xvi, 2024.
- [20] J. Ramkumar, R. Karthikeyan, and K. O. Nitish, “Securing Library Data With Blockchain Advantage,” in *Enhancing Security and Regulations in Libraries with Blockchain Technology*, 2024, pp. 117–138. doi: 10.4018/979-8-3693-9616-2.ch006.
- [21] P. S. Ponnukumar, N. I. Francis Xavier, and R. Jaganathan, “Stable Plithogenic Cubic Sets,” *J. Fuzzy Ext. Appl.*, vol. 6, no. 2, pp. 410–423, 2025, doi: 10.22105/jfea.2025.449408.1422.
- [22] V. Valarmathi and J. Ramkumar, “Modernizing Wildfire Management Through Deep Learning and IoT in Fire Ecology,” in *Machine Learning and Internet of Things in Fire Ecology*, 2024, pp. 203–229. doi: 10.4018/979-8-3693-7565-5.ch0010.
- [23] B. Suchitra, R. Karthikeyan, J. Ramkumar, and V. Valarmathi, “Enhancing Recurrent Neural Network Performance For Latent Autoimmune Diabetes Detection (LADA) Using Exocoetidae Optimization,” *J. Theor. Appl. Inf. Technol.*, vol. 103, no. 5, pp. 1645–1667, 2025.
- [24] J. Ramkumar and D. Ravindran, “Machine learning and robotics in urban traffic flow optimization with graph neural networks and reinforcement learning,” in *Machine Learning and Robotics in Urban Planning and Management*, 2025, pp. 83–104.
- [25] S. P. Priyadarshini, F. Nirmala Irudayam, and J. Ramkumar, “An Unique Overture of Plithogenic Cubic Overset, Underset and Offset,” in *Studies in Fuzziness and Soft Computing*, vol. 435, 2025, pp. 139–156.
- [26] R. Jaganathan, K. Rajendran, and P. S.

- Ponnukumar, "Peregrine Falcon Optimization Routing Protocol (PFORP) for Achieving Ultra-Low Latency and Boosted Efficiency in 6G Drone Ad-Hoc Networks (DANET)," *Int. J. Comput. Digit. Syst.*, vol. 17, no. 1, pp. 1–18, 2025, doi: 10.12785/ijcds/1571111848.
- [27] J. Ramkumar, V. Valarmathi, and R. Karthikeyan, "Optimizing Quality of Service and Energy Efficiency in Hazardous Drone Ad-Hoc Networks (DANET) Using Kingfisher Routing Protocol (KRP)," *Int. J. Eng. Trends Technol.*, vol. 73, no. 1, pp. 410–430, 2025, doi: 10.14445/22315381/IJETT-V73I1P135.
- [28] J. Ramkumar, B. Varun, V. Valarmathi, D. R. Medhunhashini, and R. Karthikeyan, "Jaguar-Based Routing Protocol (Jrp) For Improved Reliability And Reduced Packet Loss In Drone Ad-Hoc Networks (DANET)," *J. Theor. Appl. Inf. Technol.*, vol. 103, no. 2, pp. 696–713, 2025.
- [29] B. Suchitra, J. Ramkumar, and R. Karthikeyan, "Frog Leap Inspired Optimization-Based Extreme Learning Machine For Accurate Classification Of Latent Autoimmune Diabetes In Adults (LADA)," *J. Theor. Appl. Inf. Technol.*, vol. 103, no. 2, pp. 472–494, 2025.
- [30] S. P. Priyadarshini and J. Ramkumar, "Mappings Of Plithogenic Cubic Sets," *Neutrosophic Sets Syst.*, vol. 79, pp. 669–685, 2025, doi: 10.5281/zenodo.14607210.
- [31] J. Ramkumar and R. Vadivel, "CSIP—cuckoo search inspired protocol for routing in cognitive radio ad hoc networks," in *Advances in Intelligent Systems and Computing*, Springer Verlag, 2017, pp. 145–153. doi: 10.1007/978-981-10-3874-7_14.
- [32] R. Jaganathan and R. Vadivel, "Intelligent Fish Swarm Inspired Protocol (IFSIP) for Dynamic Ideal Routing in Cognitive Radio Ad-Hoc Networks," *Int. J. Comput. Digit. Syst.*, vol. 10, no. 1, pp. 1063–1074, 2021, doi: 10.12785/ijcds/100196.
- [33] K. S. J. Marseline, J. Ramkumar, and D. R. Medhunhashini, "Sophisticated Kalman Filtering-Based Neural Network for Analyzing Sentiments in Online Courses," in *Smart Innovation, Systems and Technologies*, A. K. Somani, A. Mundra, R. K. Gupta, S. Bhattacharya, and A. P. Mazumdar, Eds., Springer Science and Business Media Deutschland GmbH, 2024, pp. 345–358. doi: 10.1007/978-981-97-3690-4_26.
- [34] J. Ramkumar, A. Senthilkumar, M. Lingaraj, R. Karthikeyan, and L. Santhi, "Optimal Approach For Minimizing Delays In IoT-Based Quantum Wireless Sensor Networks Using Nn-Leach Routing Protocol," *J. Theor. Appl. Inf. Technol.*, vol. 102, no. 3, pp. 1099–1111, 2024.
- [35] J. Ramkumar and R. Vadivel, "Improved frog leap inspired protocol (IFLIP) – for routing in cognitive radio ad hoc networks (CRAHN)," *World J. Eng.*, vol. 15, no. 2, pp. 306–311, 2018, doi: 10.1108/WJE-08-2017-0260.
- [36] J. Ramkumar, S. S. Dinakaran, M. Lingaraj, S. Boopalan, and B. Narasimhan, "IoT-Based Kalman Filtering and Particle Swarm Optimization for Detecting Skin Lesion," in *Lecture Notes in Electrical Engineering*, K. Murari, S. Kamalasadan, and N. P. Padhy, Eds., Springer Science and Business Media Deutschland GmbH, 2023, pp. 17–27. doi: 10.1007/978-981-19-8353-5_2.
- [37] R. Jaganathan, S. Mehta, and R. Krishan, *Intelligent Decision Making Through Bio-Inspired Optimization*. IGI Global, 2024. doi: 10.4018/979-8-3693-2073-0.
- [38] R. Jaganathan and V. Ramasamy, "Performance modeling of bio-inspired routing protocols in Cognitive Radio Ad Hoc Network to reduce end-to-end delay," *Int. J. Intell. Eng. Syst.*, vol. 12, no. 1, pp. 221–231, 2019, doi: 10.22266/IJIES2019.0228.22.
- [39] M. P. Swapna, J. Ramkumar, and R. Karthikeyan, "Energy-Aware Reliable Routing with Blockchain Security for Heterogeneous Wireless Sensor Networks," in *Lecture Notes in Networks and Systems*, V. Goar, M. Kuri, R. Kumar, and T. Senjyu, Eds., Springer Science and Business Media Deutschland GmbH, 2025, pp. 713–723. doi: 10.1007/978-981-97-6106-7_43.
- [40] P. Menakadevi and J. Ramkumar, "Robust Optimization Based Extreme Learning Machine for Sentiment Analysis in Big Data," in *2022 International Conference on Advanced Computing Technologies and Applications, ICACTA 2022*, Institute of Electrical and Electronics Engineers Inc., 2022. doi: 10.1109/ICACTA54488.2022.9753203.
- [41] J. Ramkumar, R. Vadivel, and B. Narasimhan, "Constrained Cuckoo Search Optimization Based Protocol for Routing in Cloud Network," *Int. J. Comput. Networks Appl.*, vol. 8, no. 6, pp. 795–803, 2021, doi: 10.22247/ijcna/2021/210727.
- [42] N. K. Ojha, A. Pandita, and J. Ramkumar, "Cyber security challenges and dark side of AI: Review and current status," in *Demystifying the Dark Side of AI in Business*, IGI Global, 2024, pp. 117–137. doi: 10.4018/979-8-3693-0724-

- 3.ch007.
- [43] D. Jayaraj, J. Ramkumar, M. Lingaraj, and B. Sureshkumar, "AFSOP: Adaptive Fish Swarm Optimization-Based Routing Protocol for Mobility Enabled Wireless Sensor Network," *Int. J. Comput. Networks Appl.*, vol. 10, no. 1, pp. 119–129, 2023, doi: 10.22247/ijcna/2023/218516.
- [44] J. Ramkumar and R. Vadivel, "Improved Wolf prey inspired protocol for routing in cognitive radio Ad Hoc networks," *Int. J. Comput. Networks Appl.*, vol. 7, no. 5, pp. 126–136, 2020, doi: 10.22247/ijcna/2020/202977.
- [45] R. Vadivel and J. Ramkumar, "QoS-enabled improved cuckoo search-inspired protocol (ICSIP) for IoT-based healthcare applications," in *Incorporating the Internet of Things in Healthcare Applications and Wearable Devices*, IGI Global, 2019, pp. 109–121. doi: 10.4018/978-1-7998-1090-2.ch006.
- [46] S. P. Geetha, N. M. S. Sundari, J. Ramkumar, and R. Karthikeyan, "Energy Efficient Routing In Quantum Flying Ad Hoc Network (Q-Fanet) Using Mamdani Fuzzy Inference Enhanced Dijkstra's Algorithm (MFI-EDA)," *J. Theor. Appl. Inf. Technol.*, vol. 102, no. 9, pp. 3708–3724, 2024, [Online]. Available: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85197297302&partnerID=40&md5=72d51668bee6239f09a59d2694df67d6>
- [47] J. Ramkumar, R. Karthikeyan, and V. Valarmathi, "Alpine Swift Routing Protocol (ASRP) for Strategic Adaptive Connectivity Enhancement and Boosted Quality of Service in Drone Ad Hoc Network (DANET)," *Int. J. Comput. Networks Appl.*, vol. 11, no. 5, pp. 726–748, 2024, doi: 10.22247/ijcna/2024/45.
- [48] J. Ramkumar, R. Karthikeyan, and M. Lingaraj, "Optimizing IoT-Based Quantum Wireless Sensor Networks Using NM-TEEN Fusion of Energy Efficiency and Systematic Governance," in *Lecture Notes in Electrical Engineering*, V. Shrivastava, J. C. Bansal, and B. K. Panigrahi, Eds., Springer Science and Business Media Deutschland GmbH, 2025, pp. 141–153. doi: 10.1007/978-981-97-6710-6_12.
- [49] J. Ramkumar and R. Vadivel, "Multi-Adaptive Routing Protocol for Internet of Things based Ad-hoc Networks," *Wirel. Pers. Commun.*, vol. 120, no. 2, pp. 887–909, Apr. 2021, doi: 10.1007/s11277-021-08495-z.
- [50] R. Jaganathan, S. Mehta, and R. Krishan, *Bio-Inspired intelligence for smart decision-making*, vol. i. 2024. doi: 10.4018/9798369352762.
- [51] M. P. Swapna and J. Ramkumar, "Multiple Memory Image Instances Stratagem to Detect Fileless Malware," in *Communications in Computer and Information Science*, S. Rajagopal, K. Popat, D. Meva, and S. Bajeja, Eds., Cham: Springer Nature Switzerland, 2024, pp. 131–140. doi: 10.1007/978-3-031-59100-6_11.
- [52] J. Ramkumar, K. S. Jeen Marseline, and D. R. Medhunhashini, "Relentless Firefly Optimization-Based Routing Protocol (RFORP) for Securing Fintech Data in IoT-Based Ad-Hoc Networks," *Int. J. Comput. Networks Appl.*, vol. 10, no. 4, pp. 668–687, 2023, doi: 10.22247/ijcna/2023/223319.
- [53] M. Lingaraj, T. N. Sugumar, C. S. Felix, and J. Ramkumar, "Query aware routing protocol for mobility enabled wireless sensor network," *Int. J. Comput. Networks Appl.*, vol. 8, no. 3, pp. 258–267, 2021, doi: 10.22247/ijcna/2021/209192.
- [54] J. Ramkumar, C. Kumuthini, B. Narasimhan, and S. Boopalan, "Energy Consumption Minimization in Cognitive Radio Mobile Ad-Hoc Networks using Enriched Ad-hoc On-demand Distance Vector Protocol," 2022 *Int. Conf. Adv. Comput. Technol. Appl. ICACTA 2022*, pp. 1–6, Mar. 2022, doi: 10.1109/ICACTA54488.2022.9752899.
- [55] A. Senthilkumar, J. Ramkumar, M. Lingaraj, D. Jayaraj, and B. Sureshkumar, "Minimizing Energy Consumption in Vehicular Sensor Networks Using Relentless Particle Swarm Optimization Routing," *Int. J. Comput. Networks Appl.*, vol. 10, no. 2, pp. 217–230, 2023, doi: 10.22247/ijcna/2023/220737.
- [56] L. Mani, S. Arumugam, and R. Jaganathan, "Performance Enhancement of Wireless Sensor Network Using Feisty Particle Swarm Optimization Protocol," *ACM Int. Conf. Proceeding Ser.*, pp. 1–5, Dec. 2022, doi: 10.1145/3590837.3590907.
- [57] J. Ramkumar and R. Vadivel, "Whale optimization routing protocol for minimizing energy consumption in cognitive radio wireless sensor network," *Int. J. Comput. Networks Appl.*, vol. 8, no. 4, pp. 455–464, 2021, doi: 10.22247/ijcna/2021/209711.
- [58] T. Zhu, L. Kuang, C. Piao, J. Zeng, K. Li, and P. Georgiou, "Population-Specific Glucose Prediction in Diabetes Care With Transformer-Based Deep Learning on the Edge," *IEEE Trans. Biomed. Circuits Syst.*, vol. 18, no. 2, pp. 236–246, 2024, doi: 10.1109/TBCAS.2023.3348844.

- [59] N. N. N. Nazirun *et al.*, "Prediction Models for Type 2 Diabetes Progression: A Systematic Review," *IEEE Access*, vol. 12, pp. 161595–161619, 2024, doi: 10.1109/ACCESS.2024.3432118.
- [60] P. G. Jacobs *et al.*, "Artificial Intelligence and Machine Learning for Improving Glycemic Control in Diabetes: Best Practices, Pitfalls, and Opportunities," *IEEE Rev. Biomed. Eng.*, vol. 17, pp. 19–41, 2024, doi: 10.1109/RBME.2023.3331297.
- [61] N. F. Cleymans, M. Van De Castele, J. Vandewalle, A. K. Desouter, F. K. Gorus, and K. Barbé, "Analyzing Random Forest's Predictive Capability for Type 1 Diabetes Progression," *IEEE Open J. Instrum. Meas.*, vol. 4, pp. 1–11, 2025, doi: 10.1109/OJIM.2025.3551837.
- [62] D. Wu, Y. Mei, Z. Sun, H. Duan, and N. Deng, "Multi-Feature Map Integrated Attention Model for Early Prediction of Type 2 Diabetes Using Irregular Health Examination Records," *IEEE J. Biomed. Heal. Informatics*, vol. 28, no. 3, pp. 1656–1667, 2024, doi: 10.1109/JBHI.2023.3344765.
- [63] M. S. Al Reshan *et al.*, "An Innovative Ensemble Deep Learning Clinical Decision Support System for Diabetes Prediction," *IEEE Access*, vol. 12, pp. 106193–106210, 2024, doi: 10.1109/ACCESS.2024.3436641.
- [64] R. Jain, N. K. Tripathi, M. Pant, C. Anutariya, and C. Silpasuwanchai, "Investigating Gender and Age Variability in Diabetes Prediction: A Multi-Model Ensemble Learning Approach," *IEEE Access*, vol. 12, pp. 71535–71554, 2024, doi: 10.1109/ACCESS.2024.3402350.
- [65] Y. Sun and P. Kosmas, "Integrating Bayesian Approaches and Expert Knowledge for Forecasting Continuous Glucose Monitoring Values in Type 2 Diabetes Mellitus," *IEEE J. Biomed. Heal. Informatics*, vol. 29, no. 2, pp. 1419–1432, 2025, doi: 10.1109/JBHI.2024.3472077.
- [66] I. Shaheen, N. Javaid, N. Alrajeh, Y. Asim, and S. Aslam, "Hi-Le and HiTCL: Ensemble Learning Approaches for Early Diabetes Detection Using Deep Learning and Explainable Artificial Intelligence," *IEEE Access*, vol. 12, pp. 66516–66538, 2024, doi: 10.1109/ACCESS.2024.3398198.
- [67] T. Zhu, K. Li, P. Herrero, and P. Georgiou, "Personalized Blood Glucose Prediction for Type 1 Diabetes Using Evidential Deep Learning and Meta-Learning," *IEEE Trans. Biomed. Eng.*, vol. 70, no. 1, pp. 193–204, 2023, doi: 10.1109/TBME.2022.3187703.
- [68] S. Ahmed, M. S. Kaiser, M. S. Hossain, and K. Andersson, "A Comparative Analysis of LIME and SHAP Interpreters With Explainable ML-Based Diabetes Predictions," *IEEE Access*, vol. 13, pp. 37370–37388, 2025, doi: 10.1109/ACCESS.2024.3422319.
- [69] P. Hu *et al.*, "Prediction of New-Onset Diabetes After Pancreatectomy With Subspace Clustering Based Multi-View Feature Selection," *IEEE J. Biomed. Heal. Informatics*, vol. 27, no. 3, pp. 1588–1599, 2023, doi: 10.1109/JBHI.2022.3233402.
- [70] T. Zhu, L. Kuang, J. Daniels, P. Herrero, K. Li, and P. Georgiou, "IoMT-Enabled Real-Time Blood Glucose Prediction With Deep Learning and Edge Computing," *IEEE Internet Things J.*, vol. 10, no. 5, pp. 3706–3719, 2023, doi: 10.1109/JIOT.2022.3143375.
- [71] H. Nemat, H. Khadem, J. Elliott, and M. Benaissa, "Physical Activity Integration in Blood Glucose Level Prediction: Different Levels of Data Fusion," *IEEE J. Biomed. Heal. Informatics*, vol. 29, no. 2, pp. 1397–1408, 2025, doi: 10.1109/JBHI.2024.3483999.
- [72] A. Site, J. Nurmi, and E. S. Lohan, "Machine-Learning-Based Diabetes Prediction Using Multisensor Data," *IEEE Sens. J.*, vol. 23, no. 22, pp. 28370–28377, 2023, doi: 10.1109/JSEN.2023.3319360.
- [73] Z. Zhang *et al.*, "A novel evolutionary ensemble prediction model using harmony search and stacking for diabetes diagnosis," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 36, no. 1, p. 101873, 2024, doi: 10.1016/j.jksuci.2023.101873.
- [74] A. Sumathi, S. Meganathan, and B. V. Ravisankar, "An Intelligent Gestational Diabetes Diagnosis Model Using Deep Stacked Autoencoder," *Comput. Mater. Contin.*, vol. 69, no. 3, pp. 3109–3126, 2021, doi: 10.32604/cmc.2021.017612.