INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

ORIGINAL ARTICLE

Optimizing Gradient Boosting Machine for Alzheimer's Disease Prediction using Cormorant Swarm Intelligence (CSO-GBM)

V Haripriya[1,3]*, **P Rutravigneshwaran**[2]

1 Research Scholar, Department of Computer Science, Karpagam Academy of Higher Education, Coimbatore, Tamil Nadu, India
2 Assistant Professor, Department of Computer Science, Karpagam Academy of Higher Education, Coimbatore, Tamil Nadu, India
3 Assistant Professor, Department of Computer Technology and Data Science, Sri Krishna Arts and Science College, Coimbatore, Tamil Nadu, India

OPEN ACCESS

Received: 08/05/2025
Accepted: 20/05/2025
Published: 10/06/2025

Citation: Haripriya V, Rutravigneshwaran P (2025) Optimizing Gradient Boosting Machine for Alzheimer's Disease Prediction using Cormorant Swarm Intelligence (CSO-GBM). Indian Journal of Science and Technology 18(22): 1728-1739. https://doi.org/10.17485/IJST/v18i22.863

* **Corresponding author**.

vyharipriya.24@gmail.com

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Background: Alzheimer's Disease (AD) is a background of progressive neurodegenerative disorder which affects cognitive function and memory. In the traditional diagnostic approaches, there exists a lot of technical issues such as, hyperparameter tuning challenges, data imbalance and limited generality of predictive models result into the late detection, complex data interpretation and low classification performance. **Objectives:** The objectives are to develop an improved classification framework by enhancing the Gradient Boosting Machine (GBM) with a biologically inspired optimization algorithm and apply it to accurately and early predict AD. The desired result to accomplish this is to have a robust and adaptive learning process that can handle the above issues, including sub optimal hyperparameter selection, overfitting and sensitivity to data imbalance.**Methods:** The proposed model is Cormorant Swarm Optimized Gradient Boosting Machine (CSO-GBM) which is inspired from the cormorant's cooperative hunting and adaptive foraging behavior. The hyperparameter refinement is done efficiently by balancing the global exploration and local exploitation using the optimization algorithm. In the proposed work evaluation on the OASIS Alzheimer's MRI dataset and compared with existing models, INN-MT and FDCNN-AS. **Findings:** Results showed that the CSO-GBM model provided a high classification accuracy result of 82.39% and the highest F-measure of 83.65% as compared to benchmark models. It had shown better performance in reducing false positives and false negatives; at the same time, it accommodates high adaptability, robustness, and computational efficiency over complex data distribution. **Novelty:** This is the first research to explore swarm intelligence derived from the cormorant for optimizing GBM for predicting Alzheimer's Disease. Dynamic convergence, high generalization ability and the possibility of integration into medical diagnostic frameworks were ensured by an adaptive search mechanism; thus, the algorithm provides an excellent basis for a computationally efficient and clinically reliable classification model.

Keywords: Alzheimer's Disease Classification; Bio-Inspired Optimization; Cormorant Optimization; GBM; Feature Selection; Healthcare

1 Introduction

Alzheimer's (AD) is a progressive neurodegenerative disease that attacks memory and cognitions as well as the entire brain. This condition is characterized by the accumulation of amyloid-beta plaques and tau protein tangles that cause neural connections to deteriorate and in turn, important mental functions to be impaired. The disease seriously affects people, who have difficulties in doing their daily business, recognizing familiar faces, and speaking[1]. In advanced stages, cognitive decline occurs to a point that full dependency on caregivers is required. AD global healthcare burden is growing and calls for its early diagnosis and efficient therapies. Since early detection is important for slowing the progress of disease and increasing the patient outcomes, development of accurate and efficient diagnostic methods becomes imperative[2].

The impact of AD is not only suffering an individual, but also families, caregivers, and healthcare systems. One can expect more emotional distress, a financial burden and greater demand for long term care services as the disease progresses. The problems associated with treatment and management of AD are made more complicated due to the inability to detect it at an early stage[3]. In particular, conventional diagnostic methods, such as clinical examinations, neuroimaging and cognitive tests, commonly identify ADHD in its later stages and therefore reduce the response to therapeutic interventions. Extensive research has been done into methods for improving diagnostic accuracy that are computational ones aimed at creating accurate classification and prediction models[4].

Analysis of the above-mentioned biomarkers, such as genetic factors, cerebrospinal fluid composition, structural changes in the brain and cognitive performance metrics are needed to classify and predict AD. Machine learning (ML) and deep learning (DL) models-based development has made it possible to automate AD classification through processing of large datasets and detection of patterns associated with the disease[5]. But these models then incorporate supplemental data sources such as magnetic resonance imaging (MRI), positron emission tomography (PET) scan, and neuropsychological test result in order to increase prediction accuracy. Refining these models is essential role played by the feature selection and dimensionality reduction techniques which serve to reduce the number of attributes used in the classification process to the most relevant ones[6].

Gradient Boosting Machine (GBM) plays an important role in the prediction of AD because its model performance is improved through iterative optimization of weak learners. The multiple decision trees are constructed, with the refinement of predictions taken place each stage in minimizing the residual errors. Remarkably, GBM properly addresses this issue, and weighted error corrections help increase the accuracy of the classification. Further, the availability of hyperparameter tuning will fine tune the model to better perform, thus making it robust when handling heterogeneous data sources. GBM is a useful tool for AD diagnosis due to its ability to capture complex interactions between features, allowing to early identify disease patterns and trend of disease progression[7].

The bio inspired optimization techniques provide extra hand to improve the ML and DL models based upon the natural processes such as swarm intelligence, evolutionary adaptation, and ecological interaction. Particle swarm optimization (PSO), Genetic algorithms (GA) and Ant colony optimization (ACO) are all biological inspired algorithms that are used for hyper parameters and feature selection optimization, which results in an efficiency of the model. These methods improve exploration and exploitation, thus leading models to find better convergence and better generalization. Adaptive learning is enabled by integrating bio inspired approaches, ML and DL so that the models are robust to disparate datasets and different diagnostic conditions[8]. The objective of this work is to construct an optimized ML framework that is capable of classifying and predicting AD with bio inspired optimization and Gradient Boosting Machine. Biological inspiration in optimization method such as hyperparameter tuning, feature selection, and model performance integration with the GBM method enhanced better, more reliable, and high diagnostic results. This study is an attempt to deal with the limitations of classical classification models in improving the interpretability, the efficiency and adaptability in AD prediction. The usage of computational intelligence assists in the development of decision support systems that enable early diagnosis of patients and adjustment of the personalized treatment strategy in order to enhance patient care and healthcare management, respectively.

For the detection of Alzheimer's disease across the different age groups, the "FDCNN-AS"[9] framework applies a federated approach by using a deep convolutional neural network. This method takes advantage of the federated learning mechanism to conduct collaborative training on distributed datasets on condition that sensitive patient data does not exacerbate. The architecture for deep convolution effectively gets biomarkers from a few brain scans and helps classification. The model is able to adopt to age dependent variabilities in the process of Alzheimer's progression allowing for personalizing the accuracy in diagnosis. This technique is strengthened by advanced feature extraction, decentralized learning and it can ultimately enhance early detection capabilities combined with solving confidentiality issues for medical data processing. To perform Alzheimer's disease classification, "INNMT"[10] model that consists of a neural network framework in conjunction with multi-task learning. As the model learns multiple similar tasks at once, it comes with the added generalization, and therefore the predictive accuracy is improved. It takes in a variety of features from brain scans and learns patterns that distinguish them from varying stages of a disease. This framework uses shared representations across tasks to get better early detection and progression monitoring.

1.1 Problem Statement

Although AD is still a huge global health challenge, millions of people are still suffering from it, and it can cost top caregivers and even hospitals a great deal. It is important to detect the disease early to control the progression of the disease and to benefit the outcome of the patient. However, the traditional ways of diagnosing AD via neuroimaging and cognitive assessments fall short in capturing the very first stages of AD, thereby delaying the intervention. AD classification, however, is complicated by heterogeneous nature of clinical and imaging data, where sophisticated computational techniques need to be employed to discover meaningful information indicative of the cognitive decline. However, the accuracy rate for classifying tasks is low due to hyperparameter tuning, feature selection, and other problems with imbalanced datasets that are usually experienced with traditional machine learning models. The practical implementation of deep learning approaches is even further complicated due to the great demand of computational resources and large volume of labelled data. Disease classification using GBM has shown great results, but its performance depends on tuned hyperparameters. Bio inspired optimization techniques can ramp up GBM efficiency of learning, convergence in an adaptive way, and generalization. For the successful application for CD to AD classification and prediction, an optimized framework which uses GBM, and bio inspired strategies is necessary to increase precision of AD classification and prediction on par with human clinicians, overcome existing diagnostic shortcomings and further advance computational neurology.

2 Methodology

CSO-GBM is a bio-inspired optimization framework that mimics the adaptive and the coordinated foraging behavior of cormorants for improving the GBM hyperparameter tuning. Is an algorithm that balances the exploration (learning) and exploitation using dynamic velocity updates, leader-based influence and swarm intelligence. CSO-GBM uses iterative hyperparameter refinement to guarantee the performance of the best model, at the same time keeping diversity of the search space. The adaptive learning strategies that are integrated into the CSO-GBM lead to convergence to optimal configurations, thus improving the accuracy and efficiency of GBM related machine learning problems.

2.1 Initialize Population

Population initialization is a basic operation in utilizing the Cormorant Swarm Optimization (CSO) algorithm to optimize the Gradient Boosting Machines (GBM) such that optimal exploration of search space is guaranteed for GBMs. Once, for initialization, it involves creating a population of candidate solutions, each being a set of hyperparameters of GBM. Each of these candidates is placed inside of a specified hyperparameter space, using cormorant's natural adaptability. Specific constraints and some amount of randomization is used to determine each individual's position to allow for some amount of diverse exploration initially.

The initial population distribution within the search space uses the concept of a random uniform distribution. The position of each individual in the swarm is mathematically expressed as **Eq. (1)**.

$$X_{i,j} = L_j + r \times (U_j - L_j) \tag{1}$$

The variable $X_{i,j}$ represents the initial position of the i^{th} individual in the j^{th} dimension. L_j and U_j denote the lower and upper bounds for the j^{th} hyperparameter. r is a randomly generated number in the range [0, 1], ensuring uniform distribution across the dimensions.

The velocity initialization sets the pace of the swarm's exploration which is mathematically represented in **Eq. (2)**.

$$V_{i,j} = \sigma \times (U_j - L_j) \times \cos(\theta) \tag{2}$$

where, $V_{i,j}$ denotes the initial velocity of the i^{th} individual in the j^{th} dimension. The constant σ scales the velocity proportionally to the range of the search space. The term $\cos(\theta)$ incorporates a directional component, where θ is randomly chosen within $[0, 2\pi]$, promoting multidirectional exploration.

The behaviour of the swarm is influenced by an adaptive scaling factor during initialization, given as **Eq.(3)**.

$$\alpha = 1 - \frac{k}{K} \tag{3}$$

The variable α reduces linearly with iteration k, ensuring that initial positions prioritize exploration while subsequent updates refine exploitation. K is the total number of iterations planned for optimization.

2.2 Evaluate Initial Fitness

Evaluating the initial fitness of the swarm is essential to measure the quality of the initial positions and velocities which is assigned to each individual. Fitness evaluation directly connects to the optimization objective, where the performance of GBM is assessed based on specific hyperparameter configurations. This process ensures the swarm is guided towards promising solutions from the outset.

Each individual's fitness is calculated based on the loss function of GBM for the given hyperparameters. This is represented mathematically as **Eq.(4)**.

$$F_i = \frac{1}{M}\sum_{m=1}^{M} l\left(y_m, \hat{y}_m\left(X_i\right)\right) \tag{4}$$

where F_i represents the fitness of the i^{th} individual. y_m denotes the actual value, and $\hat{y}_m\left(X_i\right)$ is the prediction obtained using the hyperparameters encoded in X_i. The variable M is the number of data points in the validation set, and l is the loss function, such as logarithmic loss or mean squared error.

To refine the initial fitness, a gradient-based adjustment is applied as represented mathematically in **Eq.(5)**.

$$G_i = F_i - \eta.\frac{\partial F_i}{\partial X_i} \tag{5}$$

G_i represents the adjusted fitness. The term η is the learning rate for the adjustment and $\frac{\partial F_i}{\partial X_i}$ is the gradient of the fitness concerning the hyperparameters. This adjustment helps account for the local impact of changes in the hyperparameter values.

Fitness values are normalized to ensure fair comparisons across individuals which is expressed mathematically in **Eq.(6)**.

$$\widetilde{F}_i = \frac{F_i - F_{\min}}{F_{\max} - F_{\min}} \tag{6}$$

$\widetilde{F}_i$ represents the normalized fitness of the i^{th} individual. $F_{\min}$ and $F_{\max}$ are the minimum and maximum fitness values within the swarm, respectively. This ensures all fitness values are scaled between 0 and 1.

To incorporate the influence of multiple objectives, a weighted fitness contribution is defined as **Eq.(7)**.

$$W_i = \alpha.\widetilde{F}_i + (1-\alpha).\rho_i \tag{7}$$

where W_i is the weighted fitness for the i^{th} individual. The variable α is a weighting factor balancing the normalized fitness $\widetilde{F}_i$ and ρ_i, which represents a penalty term for constraints or diversity.

2.3 Identify the Global Leader

Identifying the global leader in CSO is pivotal in guiding the swarm toward optimal solutions. The global leader is determined by evaluating the fitness of all individuals and selecting the one with the best performance. This process ensures that the most promising candidate solution influences the swarm's movements, enhancing the convergence efficiency of the GBM optimization.

The global leader is identified based on the highest fitness score as represented mathematically in **Eq.(8)**.

$$L_g = argmax_{i\in\{1,2,\ldots,P\}}F_i \tag{8}$$

where, L_g represents the global leader and F_i denotes the fitness score of the i^{th} individual. P is the total population size. This equation ensures that the individual with the best performance is selected as the leader.

To ensure robustness in leader selection, a fitness stability factor is introduced as expressed in **Eq.(9)**.

$$S_g = \frac{\sum_{k=1}^{T} F_{L_g}^{k}}{T} \tag{9}$$

The term S_g represents the stability of the leader's fitness across T previous iterations. $F_{L_g}^{k}$ is the fitness of the global leader at iteration k. High stability reinforces the reliability of the chosen leader.

2.4 Update Velocity

The velocity update process in CSO governs the movement of individuals within the search space, drawing inspiration from natural adaptive behaviours. Updating velocity ensures the swarm balances exploration of new regions and exploitation around promising solutions. This step incorporates cognitive, social, and environmental influences, allowing for an optimized traversal of the hyperparameter space for GBM.

The velocity of each individual is updated as shown in **Eq.(10)**.

$$V_{i,j}^{(k+1)} = \omega . V_{i,j}^{(k)} + c_1 . r_1 . \left(P_{i,j} - X_{i,j}^{(k)}\right) + c_2 . r_2 . \left(L_{g,j} - X_{i,j}^{(k)}\right) \tag{10}$$

where, $V_{i,j}^{(k+1)}$ represents the updated velocity of the i^{th} individual in the j^{th} dimension for the $(k+1)^{th}$ iteration. ω is the inertia weight, $P_{i,j}$ is the personal best position and $L_{g,j}$ is the global leader's position. The random factors r_1 and r_2(uniformly distributed in [0, 1]) introduce stochasticity.

The leader's influence is adaptively incorporated into velocity updates which is expressed in **Eq.(11)**.

$$\Delta V_{i,j} = \frac{\lambda . \left(L_{g,j} - X_{i,j}^{(k)}\right)}{1 + e^{-\beta . C_g}} \tag{11}$$

$\Delta V_{i,j}$ denotes the additional velocity contribution from the leader, λ is an influence scaling factor, and C_g is the leader's confidence metric. The term β modulates the sensitivity to confidence.

To mimic cormorant adaptability to environmental changes, an environmental component is introduced as expressed in **Eq.(12)**.

$$V_{i,j}^{env} = \zeta . \sin(\theta) . \left(U_j - L_j\right) \tag{12}$$

The velocity $V_{i,j}^{env}$ adds an exploration term based on the range of the j^{th} dimension $\left(U_j \wedge L_j\right)$. ζ is the environmental adaptability factor, and θ is a random angle between 0 and 2π.

To prevent excessive movements, velocity is clamped using as expressed in **Eq.(13)**.

$$V_{i,j}^{(k+1)} = \min\left(\max\left(V_{i,j}^{(k+1)}, -V_{\max}\right), V_{\max}\right) \tag{13}$$

where, $V_{\max}$ is the maximum allowable velocity, limiting positional shifts to maintain stability in the optimization process.

2.5 Update Position

Updating positions in CSO is a critical process that determines the trajectory of individuals as it navigate the search space. This step uses the updated velocities from the previous step to compute the new positions of individuals, ensuring efficient exploration and exploitation of the hyperparameter space for GBM optimization. The position update process incorporates additional adaptive and environmental factors to emulate the natural behaviours of cormorants, ensuring an optimized approach to searching for the global optimum.

The position of each individual is updated as shown in **Eq.(14)**.

$$X_{i,j}^{(k+1)} = X_{i,j}^{(k)} + X_{i,j}^{(k+1)} \tag{14}$$

where, $X_{i,j}^{(k+1)}$ represents the updated position of the i^{th} individual in the j^{th} dimension for the $(k+1)^{th}$ iteration. The term $X_{i,j}^{(k)}$ is the position from the k^{th} iteration, and $V_{i,j}^{(k+1)}$ is the updated velocity computed in the previous step. This equation ensures the direct influence of velocity changes on positional adjustments.

To ensure that positions remain within valid bounds, a boundary constraint adjustment is applied as shown in **Eq.(15)**.

$$X_{i,j}^{(k+1)} = \begin{cases} L_j & if\ X_{i,j}^{(k+1)} < L_j \\ U_j & if\ X_{i,j}^{(k+1)} > U_j \\ X_{i,j}^{(k+1)} & otherwise \end{cases} \tag{15}$$

L_j and U_j represent the lower and upper bounds of the j^{th} dimension. This ensures that no individual violates the predefined search space constraints.

To emulate the dynamic spatial adaptability of cormorants, an adaptive adjustment factor is represented mathematically in **Eq.(16)**.

$$X_{i,j}^{adj} = X_{i,j}^{(k+1)} + \eta.\sin\left(\frac{\pi.k}{2K}\right) \tag{16}$$

The term $X_{i,j}^{adj}$ represents the position after applying the spatial adjustment. The factor η controls the intensity of adjustments, and the sine function ensures smooth oscillatory behaviour, enhancing the search's adaptability.

2.6 Evaluate Updated Fitness

Evaluating the updated fitness of each individual in the swarm reflects the effectiveness of position and velocity adjustments. This evaluation determines the suitability of new hyperparameter configurations for GBM performance. An optimized fitness evaluation incorporates accuracy and generalization ability, ensuring balanced progress in exploration and exploitation. Updated fitness scores guide the swarm's adaptive behaviour and influence leadership transitions, aligning to optimize GBM hyperparameters effectively.

The updated fitness is calculated using the performance of the GBM model on the validation dataset as shown in **Eq.(17)**.

$$F_i^{upd} = \frac{1}{N}\sum_{n=1}^{N} l\left(y_n, \hat{y}_n\left(X_i^{(k+1)}\right)\right) \tag{17}$$

where, F_i^{upd} is the updated fitness for the i^{th} individual. The variable N represents the total number of validation samples, y_n is the actual value and $\hat{y}_n\left(X_i^{(k+1)}\right)$ is the GBM prediction based on the updated hyperparameters encoded in $X_i^{(k+1)}$. The loss function l varies based on the task (e.g., mean squared error for regression or log-loss for classification).

To standardize fitness values and enable fair comparisons, normalization is applied.

$$\tilde{F}_i^{upd} = \frac{F_i^{upd} - F_{\min}^{upd}}{F_{\max}^{upd} - F_{\min}^{upd}} \tag{18}$$

In **Eq.(18)**, $\tilde{F}_i^{upd}$ is the normalized fitness. The terms $F_{\min}^{upd}$ and $F_{\max}^{upd}$ denote the minimum and maximum fitness values across the swarm for the current iteration. Normalization ensures fitness values lie within [0, 1], facilitating unbiased swarm adaptations.

A weighted adjustment incorporates additional factors, such as diversity or constraint penalties as expressed in **Eq.(19)**.

$$F_i^{adj} = \tilde{F}_i^{upd} - \gamma.P_i^{viol} \tag{19}$$

where, F_i^{adj} represents the adjusted fitness. The penalty coefficient γ scales P_i^{viol}, which is the magnitude of constraint violations by the i^{th} individual. This adjustment prioritizes feasible solutions during optimization.

A gradient-based refinement improves fitness evaluation accuracy by incorporating sensitivity analysis as shown in **Eq.(20)**.

$$G_i^{upd} = F_i^{adj} - \eta.\left\|\nabla F_i^{adj}\right\| \tag{20}$$

The term G_i^{upd} is the refined fitness score, where ∇F_i^{adj} represents the gradient of the fitness concerning the updated position. The parameter η controls the adjustment intensity, enhancing precision in fitness calculations.

2.7 Update Personal and Global Bests

Updating personal and global bests is crucial in CSO. This step ensures that the swarm effectively tracks progress toward optimal solutions by maintaining the best historical positions for each individual and the swarm. These updates are vital for directing the search for space exploration while balancing the exploitation of promising regions. The behaviour emulates adaptive learning, where individuals refine their search patterns by referencing their achievements and the global leader's influence.

Each updates its personal best position based on fitness improvement as expressed mathematically in **Eq.(21)**.

$$P_i^{(k+1)} = \begin{cases} X_i^{(k+1)}\ if\ F_i^{(k+1)} > F_i^{pb} \\ P_i^{(k)}\ otherwise \end{cases} \tag{21}$$

where, $P_i^{(k+1)}$ is the updated personal best position for the i^{th} individual. $F_i^{(k+1)}$ represents the fitness at the current position, and F_i^{pb} is the fitness of the personal best. This update ensures that only improvements in fitness overwrite the existing best position.

The fitness corresponding to the personal best is also updated as expressed in **Eq.(22)**.

$$F_i^{pb} = \max\left(F_i^{pb}, F_i^{(k+1)}\right) \tag{22}$$

The updated F_i^{pb} stores the maximum fitness achieved by the i^{th} individual across iterations, aligning with its updated personal best position.

The global best position is updated by evaluating the fitness of all individuals is represented mathematically in **Eq.(23)**.

$$G^{(k+1)} = argmax_{i\in\{1,2,\ldots,P\}} F_i^{pb} \tag{23}$$

where, $G^{(k+1)}$ is the global best position for the $(k+1)^{th}$ iteration, determined by the personal bests of all individuals. This ensures the swarm aligns its movements with the most optimal solution.

A stability metric is introduced to assess the consistency of the global best leader as defined mathematically in **Eq.(24)**.

$$S_g^{(k+1)} = \frac{\sum_{t=k-m}^{k} F_G^{(t)}}{m} \tag{24}$$

The metric $S_g^{(k+1)}$ averages the fitness of the global best across the last m iterations, providing a measure of the leader's robustness over time.

2.8 Repeat Optimization Cycle

The optimization cycle in CSO iteratively refines the search space by repeatedly executing a structured set of operations. This cyclical process ensures continuous improvement of the swarm's performance, leveraging adaptive behaviours to converge toward the optimal hyperparameters for GBM tuning. Repeating the cycle allows the swarm to balance exploration and exploitation dynamically, adjusting its behaviour to suit the characteristics of the search space.

Each cycle begins with updated velocity and position calculations, incorporating swarm dynamics as shown in **Eq.(25)**.

$$\begin{gathered} V_{i,j}^{(k+1)} \\ = \\ \omega^{(k)}.V_{i,j}^{(k)} + c_1.r_1.\left(P_{i,j} - X_{i,j}^{(k)}\right) + c_2.r_2.\left(G_j - X_{i,j}^{(k)}\right) \\ X_{i,j}^{(k+1)} \\ = \\ X_{i,j}^{(k)} + V_{i,j}^{(k+1)} + \xi.\sin(\theta) \end{gathered} \tag{25}$$

where, $V_{i,j}^{(k+1)}$ and $X_{i,j}^{(k+1)}$ denote the updated velocity and position of the i^{th} individual in the j^{th} dimension. The velocity incorporates personal best $P_{i,j}$, global best G_j, and adaptive components such as random influences r_1, r_2, and environmental oscillations $\xi \cdot \sin(\theta)$.

Updated positions are evaluated for fitness using a multi-objective approach as expressed in **Eq.(26), Eq.(27)** and **Eq.(28)**.

$$F_i^{(k+1)} = \alpha.Acc_{GBM}^{(k+1)} - \beta.Comp_{GBM}^{(k+1)} \tag{26}$$

$$Acc_{GBM}^{(k+1)} = \frac{1}{N}\sum_{n=1}^{N} l\left(y_n, \hat{y}_n\left(X_i^{(k+1)}\right)\right) \tag{27}$$

$$Comp_{GBM}^{(k+1)} = \frac{Time_{train} + Time_{eval}}{2} \tag{28}$$

The fitness $F_i^{(k+1)}$ combines accuracy $Acc_{GBM}^{(k+1)}$ and computational cost $Comp_{GBM}^{(k+1)}$, where α and β balance these objectives. Training and evaluation times reflect the computational overhead of GBM.

Cycle iterations adapt key swarm parameters dynamically which is expressed using **Eq.(29)** and **Eq.(30)**.

$$\omega^{(k+1)} = \omega_{\max} - \frac{k}{K}.(\omega_{\max} - \omega_{\min}) \tag{29}$$

$$c_1^{(k+1)} = c_1^{(k)} + \delta.\sin\left(\frac{\pi k}{2K}\right) \tag{30}$$

The inertia weight $\omega^{(k+1)}$ decreases linearly over iterations, promoting convergence. The cognitive coefficient $c_1^{(k+1)}$ oscillates with a sinusoidal adjustment $\delta.\sin\left(\frac{\pi k}{2K}\right)$, enhancing exploration during early cycles.

Variance metrics assess swarm diversity are expressed in **Eq.(31)** and convergence in **Eq.(32)**.

$$\sigma_X^{(k+1)} = \sqrt{\frac{1}{P}\sum_{i=1}^{P}\left\|X_i^{(k+1)} - \overline{X}^{(k+1)}\right\|^2} \tag{31}$$

$$\overline{X}^{(k+1)} = \frac{1}{P}\sum_{i=1}^{P} X_i^{(k+1)} \tag{32}$$

Positional variance $\sigma_X^{(k+1)}$ measures the spread of the swarm, while $\overline{X}^{(k+1)}$ represents the centroid of updated positions. These metrics guide adjustments to enhance exploration if variance falls below a threshold.

Constraint handling ensures valid solutions as shown in **Eq.(33)**.

$$X_{i,j}^{(k+1)} = \begin{cases} U_j \ if \ X_{i,j}^{(k+1)} > U_j \\ L_j \ if \ X_{i,j}^{(k+1)} < L_j \\ X_{i,j}^{(k+1)} \ otherwise \end{cases} \tag{33}$$

Positional corrections maintain feasibility by enforcing bounds $\left[L_j, U_j\right]$ for each dimension. This reinforcement aligns the swarm's trajectory with problem constraints.

The global leader's influence evolves adaptively.

$$L_g^{(k+1)} = argmax_{i\in\{1,2,\ldots.P\}} F_i^{pb} \tag{34}$$

$$R_g^{(k+1)} = R_g^{(k)}.\left(1 - \frac{\sigma_X^{(k+1)}}{\sigma_X^{(0)}}\right) \tag{35}$$

In **Eq.(34)**, leader $L_g^{(k+1)}$ reflects the individual with the highest personal fitness. Its influence radius $R_g^{(k+1)}$ is shown in **Eq.(35)**, narrows with decreasing positional variance $\sigma_X^{(k+1)}$, enhancing exploitation.

Progress toward convergence is monitored using fitness variance are expressed using **Eq.(36)** and **Eq.(37)**.

$$\sigma_F^{(k+1)} = \sqrt{\frac{1}{P}\sum_{i=1}^{P}\left(F_i^{(k+1)} - \overline{F}^{(k+1)}\right)^2} \tag{36}$$

$$\overline{F}^{(k+1)} = \frac{1}{P}\sum_{i=1}^{P} F_i^{(k+1)} \tag{37}$$

Fitness variance $\sigma_F^{(k+1)}$ and mean fitness $\overline{F}^{(k+1)}$ quantify swarm progress. Low variance signals convergence, guiding termination decisions.

2.9 Extract Optimal Hyperparameters

Extracting the optimal hyperparameters is the culmination of the CSO process. This step identifies the most effective configuration for the GBM model from the swarm's search. Optimal hyperparameters are determined by evaluating the personal and global bests while ensuring alignment with the optimization goals. This process reflects the swarm's adaptability and ability to converge toward the most promising solution in the hyperparameter space.

The optimal position in the search space is identified as the global best at the final iteration is expressed in **Eq.(38)**.

$$X_{opt} = G^{(K)} \tag{38}$$

where, X_{opt} represents the optimal hyperparameter configuration and $G^{(K)}$ is the global best position at the final iteration K. This position reflects the most effective solution identified by the swarm throughout the optimization process.

The fitness corresponding to the optimal configuration is evaluated to validate its effectiveness as shown in **Eq.(39)**.

$$F_{opt} = \frac{1}{N}\sum_{n=1}^{N} l\left(y_n, \hat{y}_n\left(X_{opt}\right)\right) \tag{39}$$

where, F_{opt} denotes the fitness of the optimal configuration. The term y_n represents the actual target values, $\hat{y}_n\left(X_{opt}\right)$ is the prediction from the GBM model using X_{opt}, and l is the loss function corresponding to the optimization objective.

The stability of the optimal solution is assessed using positional variance across recent iterations as shown in **Eq.(40)**.

$$\sigma_{opt}^2 = \frac{1}{m}\sum_{k=K-m+1}^{K} \left\|G^{(k)} - G^{(k)}\right\|^2 \tag{40}$$

The metric σ_{opt}^2 evaluates the consistency of the global best position $G^{(k)}$ over the last m iterations. Low variance indicates the stability and reliability of the extracted solution.

The robustness of the optimal configuration is analyzed by perturbing the solution slightly and observing performance as expressed in **Eq.(41)**.

$$R_{opt} = \frac{1}{P}\sum_{i=1}^{P} l\left(y_n, \hat{y}_n\left(X_{opt} + \Delta X_i\right)\right) \tag{41}$$

The robustness score R_{opt} averages the loss l over P perturbations ΔX_i. This ensures that the solution is not overly sensitive to small changes in hyperparameters, confirming its generalization ability.

A refinement process adjusts the optimal solution for better accuracy as expressed in **Eq.(42)**.

$$X_{refined} = X_{opt} - \eta.\nabla F_{opt} \tag{42}$$

The refined solution $X_{refined}$ is derived by moving in the direction of the negative gradient -∇F_{opt} of the fitness function. The step size η controls the intensity of this adjustment, ensuring precision without significant deviations.

3 Results and Discussion

3.1 About Dataset and Metrics

The OASIS Alzheimer's Detection Dataset is a foundational resource for conducting research on early stages of neurodegeneration. It consists of 1012 MRI images, where 796 are used for training and 216 for testing, that are well structured for machine learning applications. The dataset is comprised of T1-weighted MRI scans to visualize structures of the brain related to Alzheimer's progression in great detail. It consists of five categories of Normal Control (NC), Mild Cognitive Impairment (MCI), and Alzheimer's Disease (AD), which are able to accurately classify. The other biomarkers consist of important attributes such as age, gender and hippocampal volume that could further improve predictive modeling. It possesses features for the extraction of key structural features – namely the cortical thickness, ventricular enlargement, and brain atrophy patterns, which support computer aided diagnosis (CAD) systems. As it is structured organization it suits the purpose for deep learning architectures and boosting up the accuracy of predictive models of early Alzheimer's detection and prognosis.

In this section, classification accuracy and the F measure are used to compare the efficiency of FDCNN, AS, INN, MT and CSO, GBM. Finally, these metrics offer a view of how well the models in achieving true positives, true negatives and false minimal classifications. Accuracy of classification is the number of correctly classified instance divided by the total number of instances. A higher classification accuracy would be better performing in the sense of distinguishing categories. According to the resultant outcomes, CSO-GBM has been able to yield the highest accuracy classification, that is, 82.3940%, outperforming FDCNN-AS, which is 53.4830% and INN-MT, 57.0970%. This is a significant improvement and underscores that the CSO – GBM framework is capable to dealing with complex data distributions as well as optimize the classification processes. The reason why CSO-GBM has the superior accuracy is that it uses bio-inspired optimization to improve feature selection and search of hyperparameters in order to achieve the best performance. FDCNN-AS and INN-MT have relatively poor classification accuracy, which translates to their incapability to capture intricate patterns of the dataset. The values for classification accuracy provided above, when visualized in Figure 1, reinforce how much CSO-GBM has improved upon the accuracy results of its predecessor.

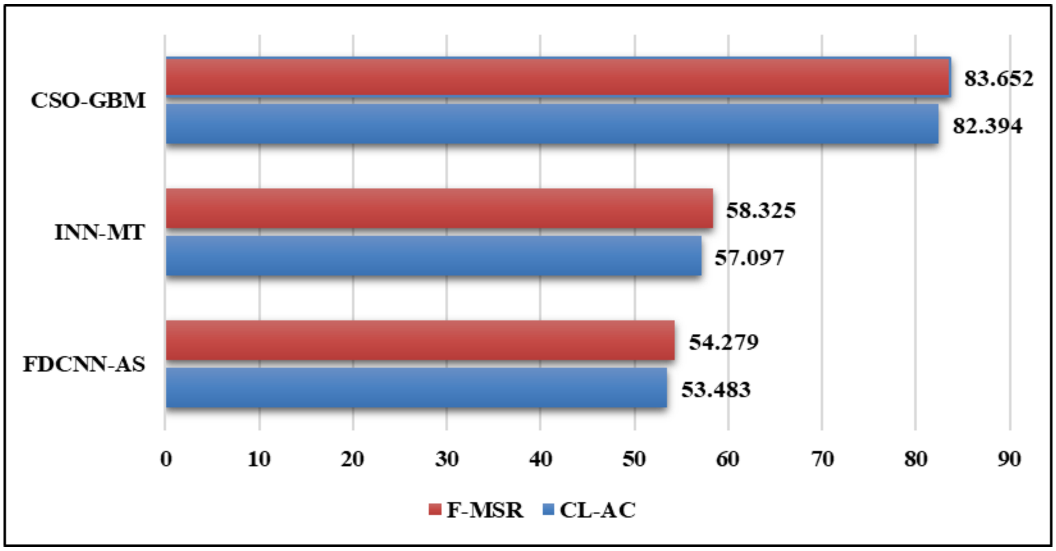

Fig 1. Classification Accuracy and F-Measure

A type of harmonic mean of the precision and recall is the F-measure (called F1-score) which gives a good measure of a model's performance in separating positive instances from the negative ones while minimizing the false positives. A higher F-measure is a better trade-off between precision and recall. This provides CSO-GBM model with an F-measure of 83.6521 featuring a considerable predictive reliability improvement. F-measure value of 58.3252 and 54.2790 is registered in FDCNN-AS and INN-MT, respectively. Improvement of this CSO-GBM provides an indication that the model retains high classification accuracy and minimizes misclassification by differentiating between exactness and inexactness of actual and random classifications.

The Fowlkes–Mallows index (FMI) is an indication of how well a classification model separates between positive and negative cases, being the geometric mean of precision and recall. A higher FMI value means that predictions are more related to actual classifications. Results show that the CSO-GBM has the highest FMI of 83.719 which is a very large improvement over the INN-MT which scored 58.461 and FDCNN-AS which scored 54.405. Their enhanced performance on the CSO-GBM shows that it offers a balanced (strong) trade-off between precision and recall, thus making it suited for accurately classifying instances.

Comparison of the models using the parameters FMI and MCC is presented in Figure 2. A robust quality metric for binary classifications is used, it considers all confusion matrix elements true positives, true negatives, false positives and false negatives, the Matthews Correlation Coefficient (MCC). For higher values of MCC, it means that predictive accuracy as well as overall classification reliability is strong. The MCCs of CSO-GBM is 64.888, while INN-MT obtained an MCC of 14.686 and FDCNN-AS recorded 7.452. These indicate that CSO-GBM has much better ability to make balanced and accurate predictions among different categories of classifications.

The precision (PREC) is retrieved by the proportion of the instances that get correctly classified as a positive instance over the total number of positive instances classified. A higher value for the precision score would imply that a model was good at reducing the number of false positives without reducing the accuracy of the positive classifications. For example, CSO-GBM has

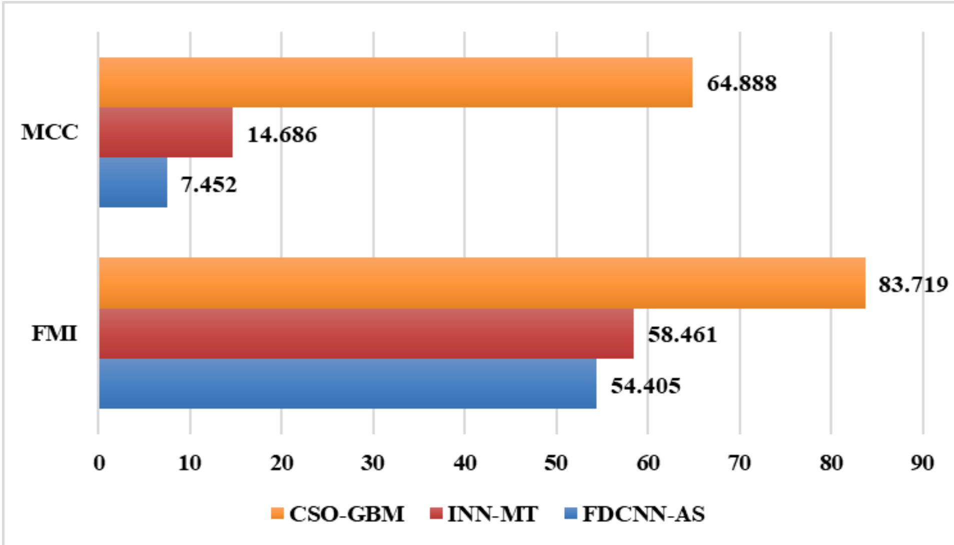

Fig 2. FMI and MCC

gained the highest precision of 80.430 that increases considerably compared to INN-MT and FDCNN-AS that are 54.600 and 50.825, respectively. By the higher precision of CSO—GBM, it is shown to be able to optimize the classification boundaries with well-tuned hyper parameters which in turn, enhances the actual positives and misclassified positives discrimination. Another traditional model, FDCNN-AS and INN-MT, have worse precision values meaning that they are more likely to wrongly classify an RPG to the positive class.

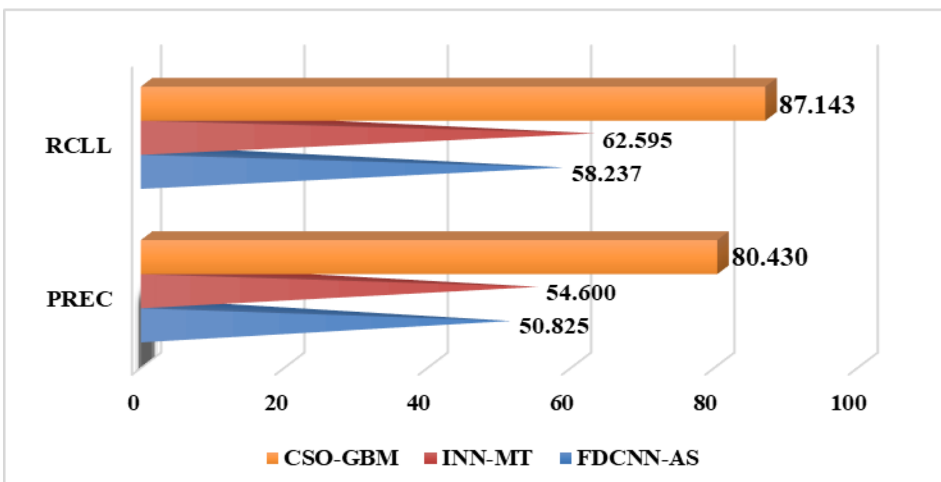

Fig 3. PREC and RCLL

And the other is Recall (RCLL) i.e., True Positive Rate, which is the proportion of the actual positive instances classified correctly by a model. A higher recall value indicates that the model has a better capacity to attain true positive cases while reducing false negatives. It is seen that CSO-GBM has significant recall improvement (87.143) compared to the others (62.595 and 58.237 for INN-MT and FDCNN-AS respectively). The improvement in CSO-GBM for recall is substantial and implies that the model accurately identifies positive instances without classifying them as negatives. Recall values of the traditional models are lower, implying that those models cannot find all relevant cases because of lack of learning adaptation.

4 Conclusion

The combination of cormorant swarm intelligence with Gradient Boosting Machine has developed a new approach for classifying Alzheimer's Disease. Using the CSO-GBM framework has led to better classification performance, balanced accuracy and recall, and better generalization compared with common models. Adaptive hyperparameter tuning and managing features effectively are made possible with biologically inspired search algorithms and help deal with major challenges such as overfitting, imbalance, and inefficient converging in datasets used in clinical study. Instead of using standard methods, this work applies the foraging ideas of cormorants to the optimization of features in a non-linear setting. As a result of the balance between exploration and exploitation, CSO-GBM is appropriate for early AD prediction and can grow steadily. It is also on the way to being used in multi-modal classification, prediction over time, and linking to electronic patient data. It is possible that in the future, swarm intelligence will be used in real-time diagnostic systems for various chronic neurological problems. It would be useful if future research looked into combining models, using bigger data sets, and applying these models to practical systems for use by clinicians.

References

1) Thulasimani V, Shanmugavadivel K, Cho J, Easwaramoorthy SV. A Review of Datasets, Optimization Strategies, and Learning Algorithms for Analyzing Alzheimer's Dementia Detection. *Neuropsychiatric Disease and Treatment*. 2024;20:2203–2225. https://doi.org/10.2147/NDT.S496307.
2) Chen Y, et al. Automated Alzheimer's Disease Classification Using Deep Learning Models with Soft-NMS and Improved ResNet50 Integration. *Journal of Radiation Research and Applied Sciences*. 2024;17(1):100782. https://doi.org/10.1016/j.jrras.2023.100782.
3) Dubey Y, Bhongade A, Palsodkar P, Fulzele P. Efficient Explainable Models for Alzheimer's Disease Classification with Feature Selection and Data Balancing Approach Using Ensemble Learning. *Diagnostics*. 2024;14:2770. https://doi.org/10.3390/diagnostics14242770.
4) Habib M, Vicente-Palacios V, García-Sánchez P. Bio-Inspired Optimization of Feature Selection and SVM Tuning for Voice Disorders Detection. *Knowledge-Based Systems*. 2025;310. https://doi.org/10.1016/j.knosys.2024.112950.
5) Ibrahim R, Ghnemat R, Al-Haija QA. Improving Alzheimer's Disease and Brain Tumor Detection Using Deep Learning with Particle Swarm Optimization. *AI*. 2023;4(3):551–573. https://doi.org/10.3390/ai4030030.
6) Shojaie M, Cabrerizo M, DeKosky ST, Vaillancourt DE, Loewenstein D, Duara R, et al. A Transfer Learning Approach Based on Gradient Boosting Machine for Diagnosis of Alzheimer's Disease. *Frontiers in Aging Neuroscience*. 2022;p. 966883. https://doi.org/10.3389/fnagi.2022.966883.
7) Balakrishnan NB, Pillai AS, Panackal JJ, Sreeja PS. Alzheimer's Disease Detection and Classification Using Optimized Neural Network. *Computers in Biology and Medicine*. 2025;187. https://doi.org/10.1016/j.compbiomed.2025.109810.
8) Menakadevi P, Ramkumar J. Robust Optimization Based Extreme Learning Machine for Sentiment Analysis in Big Data. 2022;p. 1–5. https://doi.org/10.1109/ICACTA54488.2022.9753203.
9) Lakhan A, et al. FDCNN-AS: Federated Deep Convolutional Neural Network Alzheimer Detection Schemes for Different Age Groups. *Information Sciences (N Y)*. 2024;677:120833. https://doi.org/10.1016/j.ins.2024.120833.
10) Zhang X, Gao L, Wang Z, Yu Y, Zhang Y, Hong J. Improved Neural Network with Multi-Task Learning for Alzheimer's Disease Classification. *Heliyon*. 2024;10(4):e26405. https://doi.org/10.1016/j.heliyon.2024.e26405.