NavaJyoti, International Journal of Multi-Disciplinary Research
Volume 7, Issue 1, August 2022

Feature Selection Strategy Based on Meta Heuristics Optimization Algorithm and Chaotic Maps – A Review

M. Janaki[1], Dr. S.N.Geethalakshmi[2]
[1]*PhD Research Scholar, Dept of Computer Science, Avinashilingam University for Women, Coimbatore*
janakisanjay81@gmail.com
[2]*Professor, Dept of Computer Science, Avinashilingam University for Women, Coimbatore*
sngeethalakshmi@gmail.com

Abstract:

Feature selection is a common and an important procedure in classification problem to achieve a better accuracy and faster convergence. The purpose of this paper is to provide an introduction of chaos theory, various optimisation algorithms like particle swarm optimization, grey wolf optimization, whale optimization, genetic algorithm and review of current research papers on meta heuristic optimization algorithm. Finally it concludes existing optimisation algorithm have slow convergence speed therefore, chaos strategy used in optimization due to their dynamic behaviour which help optimization algorithm exploring the search space more dynamically and globally.

Keywords*:* optimization algorithm, chaos theory, Feature Selection, Applications

1. Introduction :

1.1. MetaHeuristics

A metaheuristic is a set of algorithmic concepts that can be used to define heuristic methods applicable to a wide set of different problem. Metaheuristic algorithms are typically nature-inspired, and that they are currently among the foremost wide used algorithms for optimization. They have several benefits over typical algorithms, simulated hardening, bee algorithms, particle swarm optimization, harmony search, firefly algorithmic rule, cuckoo search, whale optimization, and gray wolf optimization. Intensification and diversification are two fundamental aspects of metaheuristics. Diversification ensures that the algorithm examines the search space broadly, whereas intensification aims to search locally and more intensely

2. Optimization Algorithm:

Optimization is everywhere, and is thus an important paradigm itself with a wide range of applications. In almost all applications in engineering and industry, we are always trying to optimize something - whether to minimize the cost and energy consumption, or to maximize the profit, output, performance and efficiency. The optimization techniques are used to find out the best among the possible combinations of the feature subsets that helps to reduce the computational complexity. The exhaustive

search of feature subsets can be optimized by certain meta-heuristic approaches that help to find out the global best among search criteria. Here we'll introduce a number of these algorithms shortly.

2.1 Particle swarm optimization (PSO):

The idea arises from the natural behavior that a large number of birds flock parallelly, change direction spontaneously, scatter and regroup at intervals, and finally reach a target. PSO is simplicity and generality. Is inspired by social behavior of birds flocking and Swarm – population of moving particles .Each member of the population is called particle. All the particles have a fitness value. The fitness values can be calculated using objective function .All the particles preserve their individual best performance. They also know the best performance of their group. They adjust their velocity considering their best performance of the best particle.[8]
The steps of particle swarm optimization are as follows:

Step 1: Initialization – The swarm particles lie within the pre-defined ranges of velocity and position.

Step 2: Velocity Updating – At every cycle the speeds of the swarm particles are calculated by: Eq.(1)

$$\vec{V}_i = W\vec{V}_i + c_1R_1(\vec{P}_{i,best} - \vec{P}_i) + c_2R_2(\vec{g}_{i,best} - \vec{P}_i) \qquad (1)$$

Where$\vec{P}_i$ = position of particle 'i'
$\vec{V}_i$ = velocity of particle 'i'
$\vec{P}_{i,best}$ = finest position reached by the particle
$\vec{g}_{i,best}$ =best location remembered by the particle individual
'W' = parameter controlling the flying elements
R_1, R_2 =random numbers among 0 and 1
c_1, c2 =cognitive learning factor and social learning factor

Step 3:Updating of position – There is an interval among succeeding iterations and hence the positions of the particles undergo change.

$$\vec{P}_i = \vec{P}_i + \vec{V}_i \qquad (2)$$

After refreshing, $\vec{P}_i$ must be verified and in the allowable range.

Step 4: Updating of memory – Update $\vec{P}_{i,best}$ and $\vec{g}_{i,best}$ using the formula

$$\vec{P}_{i,best} = \vec{P}_i iff(\vec{P}_i) > f(\vec{P}_{i,best}) - \qquad (3)$$

$$\vec{g}_{i,best} = \vec{g}_i iff(\vec{g}_i) > f(\vec{g}_{i,best}) - \qquad (4)$$

Where $(\vec{x})$ is the point function subject to extension.

Step 5: Destination Checking – The technique iterates steps 2 to 4 until definite end states are reached, for a specified number of iterations, when ended. The estimation of $\vec{g}_{i,best}$ and $\vec{P}_{i,best}$ give the result.

The fitness values are not considered in PSO algorithms. This is a big computational advantage over GA, when the population is huge. Arithmetic operation of real numbers is used for calculation of velocity and position. The PSO works on the following principles namely modifying the

velocity (accelerating) toward pbest (Personal Best) and lbest (Global Best) locations(local version of PSO) and accelerating towards pbest and lbest locations .

Personal Best:The best solution (fitness value) is effectively achieved when individual particle maintains the path of its coordinates in feature space. This value is called pbest.

Global Best: Particle, while maintaining all the features as its topological neighbor, this value is the gbest (global best). When a particle locates with reference to its neighbor it is called lbest.

2.2 Grey wolf Optimizer Algorithm (GWO):
It is a novel type of optimization algorithm and used to update the current position of population in the discrete searching space, thus getting the optimal feature subset. It is a novel meta-heuristic, inspired from the social hunting behavior of grey wolves. It also has good convergence ability towards the optima. Four types of grey wolves such as alpha, beta, delta, omega.[9]

Main steps are hunting, searching for prey, encircling prey. Alpha – making responsible for making decision about hunting.Takes feedback from other wolves and give it to the alpha leader.

In Sankara Babuet.al [9] the author discussed to predict the different kinds of diseases using Grey Wolf optimization and auto encoder based Recurrent Neural Network (GWO+RNN) and The GWO+RNN method achieved 16.825 %of improved accuracy in Cleveland dataset for disease prediction.

In QiangLi et.al [10] the author showed a new predictive framework is proposed by integrating an improved grey wolf optimization (IGWO) and kernel extreme learning machine (KELM), termed as IGWO-KELM, for medical diagnosis and Gains less number of selected features, achieving high classification performance

2.3 Whale optimization algorithm (WOA):

A rapid way to tackle a wide range of optimization issues. It is one of the new meta-heuristic optimization algorithms, which is inspired by social hierarchy and the intelligent hunting method of grey wolves. These two social behaviors of gray wolves' pack (social hierarchy and hunting technique) are modeled in the GWO algorithm. The main idea of the WOA is inspired by the hunting behavior of the humpback whales. Whales in the ocean move in a spiraling motion towards their prey, creating bubbles to encircle it. As for any metaheuristic algorithm, WOA has two phases; exploitation and exploration.
In the following section, we will describe the mathematical model of encircling prey, searching for prey, and spiral bubble-net foraging man oeuvre.

Encircling prey: By the increasing number of iterations from start to a maximum numbers, humpback whales encircle the prey and update their position in the direction of the best search agent. We can mathematically formulate this behavior as:

If (p<0.5 and mod (U) <1)

Then the position of the candidate position X (t+1) is updated by the following equations .Eq(5)

$$D = \text{mod}\{(C.X)-X(t)\} - \quad (5)$$

$$X(t+1) = [X(t) - \{U.D\}]$$

Where p =0.1 (constant) X (t+1) is the best position in the current situation. U and D are calculated by the following equations - Eq.(6)

$$U = \text{mod}\{2.a.r-a\} - \quad (6)$$

$$C=2.r$$

Where a is linearly decreases from 2 to 0 and r is the randomly selected vector

Prey Searching: In prey searching mechanism, X is replaced with the random variables X_{random} and mathematical equation are given as follows

$$D = \text{mod}\{(C.X_{random})-X(t)\} - \quad (7)$$

$$X(t+1) = [X_{random}(t) - \{U.D\}] - \quad (8)$$

The encircling the prey and spiral updating of the prey has been done during the exploration phase of whale optimization algorithm. The mathematical expression for updation of new position during the spiral process is given by below equation

$$X(t+1) = D^{l}.e^{bl}.\cos(2\pi l) + X^{*}(t) - \quad (9)$$

Where D is the distance between the new position and updated position in new generation, b is the constant which varies from the 0 to 1.

Yancang Li et al., proposed a novel modified ***whale optimization*** algorithm based on Tent chaos map and tournament selection strategy (MWOA) that improved algorithm is utilized to reduce the possibility of the standard whale algorithm falling into local optimal. During the initialization of the population, in order to increase population diversity and randomness, MWOA cites the Tent chaos map. In order to improve the development ability of the optimization algorithm, the tournament selection strategy was employed to improve the algorithm accuracy..

2.4 Genetic Algorithm(GA):

A genetic Algorithm is a population genetics-based adaptive heuristic search strategy. It's a probabilistic search approach based on the natural selection and reproduction process. It's utilized to find answers to optimization issues. It is an evolutionary algorithm that combines initialization, selection, crossover, and mutation. Selection process is a process by which the number of features is reduced into a compact subset subsequently bringing high accuracy in classification. In Fig 1 the genetic algorithm process is shown completely.

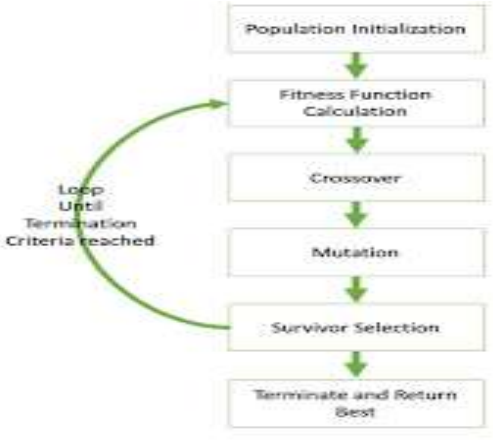

Figure 1: *Genetic Algorithm*

2.4.1Applications of Genetic Algorithm

Genetic algorithms have been used for many real life problems. Some applications of GA are Combinatorial Optimization, traveling salesman (TSP), Sequence scheduling, routing, bin packing, graph coloring and partitioning, nonlinear dynamical systems predicting, data analysis, functions for creating images and strategy planning .

2.5 Ant Colony Method(ACO):

Ant colony optimization (ACO) takes inspiration from the foraging behavior of some ant species. These ants deposit pheromone on the ground in order to mark some favorable path that should be followed by other members of the colony. Ant colony optimization exploits a similar mechanism for solving optimization problems.[11]

3. Feature selection:

Feature selection is the method of picking out the important and optimal features from the set of high dimensional features. It is used to optimize the feature set in the extracted features, where the number of features was reduced by removing irrelevant and redundant feature to achieve acceptable classification accuracy. [19].Feature selection is to be converted to a more reliable and suitable form of input for the classifier to classify the different features category.

Dataset is available in UCI repository and dataset is the core for every predictive model. Each dataset is normalized, and for evaluation, 80% of data were used for training and 20% of data were used for testing. The evaluation is carried out with the accuracy parameter. The training set is utilized for the

known data process, while the test set is applied for unknown data . In fig 2 the flow chart diagram of feature selection are explain in detail.

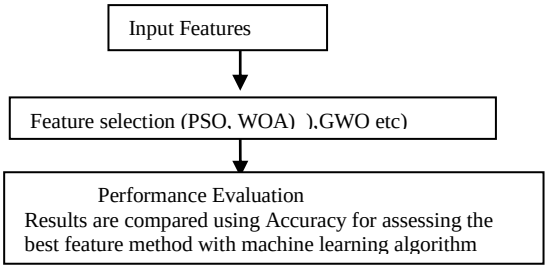

Figure 2: *Flow Diagram of Feature Selection*

4. Literature Review:

Longzhen Duan .(2020) utilized a new feature selection strategy for niche PSO algorithm based on the chaos group, which is used to assess the significance of feature selection techniques The experimental findings reveal that the algorithm's feature selection performance is superior to that of the comparison method, and that classification accuracy has increased greatly.[12]

Arshad et al.(2020) examines that this study provides a comprehensive overview of chaos theory and how it might be used to visual encryption techniques. Chaos-based image encryption systems offer various benefits over regular encryption techniques due to ergodicity and initial key sensitivity. The paper explores some of the most important applications of chaos theory, notably in the field of image encryption.[13]

Yancang Li et al.(2020), proposed a novel modified ***whale optimization*** algorithm based on Tent chaos map and tournament selection strategy (MWOA) that improved algorithm is utilized to reduce the possibility of the standard whale algorithm falling into local optimal. During the initialization of the population, in order to increase population diversity and randomness, MWOA cites the Tent chaos map. In order to improve the development ability of the optimization algorithm, the tournament selection strategy was employed to improve the algorithm accuracy.[14]

Yintong Li .(2019) proposed whale optimization algorithm with ***chaos strategy*** and weight factor (WOACW) method in which, chaos strategy used for initialize the population to enhance the diversity of the initial population. The weight factor is introduced to adjust the influence degree of the current optimal solution on the generation of new individuals in order to improve the convergence speed and accuracy. At the same time, the convergence factor is adjusted by cosine function to better balance the relationship between exploration and exploitation. Also, using greedy strategy to fully retain the dominant individuals from the parents and the generated candidates to generate offspring, improves the convergence speed of the proposed (WOACW) algorithm.[15]

M.Kohli,S.Arora.(2018) examines that Grey Wolf Optimization (GWO) algorithm is a new meta-heuristic that is based on grey wolf social hunting behavior. The goal of this study is to include chaos theory into the GWO algorithm in order to improve its global convergence speed. The findings indicated

that, given the right chaotic map, CGWO may clearly outperform traditional GWO, with excellent results in contrast to other methods and in limited optimization situations.[16]

Fei Ye.(2017) emphasized that this work provides a novel SVM optimization approach based on an enhanced chaotic fly optimization algorithm (FOA) with a mutation strategy to accomplish parameter setup turning for the SVM and feature selection at the same time. It has been demonstrated that it is a more resilient and successful optimization strategy than other well-known methods, especially when tackling medical diagnosis and credit card problems.[17].

In table 1 the current papers on feature optimization techniques are summarized in detail manner .

Studies	Paper Title & years	Authors	Concepts discussed	Conclusion	Feature optimization techniques
[1]	Heart Disease Prediction and Classification Using Machine Learning Algorithms Optimized by Particle Swarm Optimization and Ant Colony Optimization (2019)	Youness Khourdifi1* Mohamed Bahaj1	Study examines the different machine learning algorithms and compares the results using different performance measures, i.e. accuracy, precision, recall, f1-score	Experimental results show that the optimization hybrid approach increase the predictive accuracy of medical data sets.	Particle Swarm Optimization and Ant Colony Optimization
[2]	A whale optimization algorithm (WOA) approach for clustering (2018)	Jhila Nasiri1 and Farzin Modarres Khiyabani1*	The whale optimization algorithm developed to solve popular clustering problem	Results of this algorithm is compared with well-known k-means clustering approach and other popular stochastic algorithms such as PSO, ABCand GA.	Whale optimization algorithm
[3]	Whale Optimisation Algorithm for high dimensional small-instance feature selection (2019)	Majdi Mafarja, Iyad Jaber, Sobhi Ahmed & Thaer Thaher	The challenge in this paper was the use of high-dimensional datasets, with a low number of samples	WOA approach proved a superior performance on three evaluation criteria; classification accuracy, the best fitness value, and the number of selected features.	Whale Optimisation Algorithm
[4]	Whale Optimization Algorithm With Applications to Resource Allocation in Wireless Networks(2020)	Quoc-VietPham ,Member,IEEE,	To study the applicability of WOA to solve resource allocation problems in wireless networks.	To demonstrate the applicability of WOA in wireless and communication networks.	Whale Optimization Algorithm
[5]	Medical Disease Prediction using Grey Wolf optimization and Auto Encoder based Recurrent Neural Network (2018)	B Sankara Babu1, A Suneetha, G Charles Babu, Y Jeevan	To predict the different kinds of diseases using Grey Wolf optimization and auto encoder based Recurrent Neural Network (GWO+RNN).	The GWO+RNN method achieved 16.825% of improved accuracy in Cleveland dataset for disease prediction.	Grey Wolf optimization
[6]	An Enhanced Grey Wolf Optimization Based Feature Selection Wrapped Kernel Extreme Learning Machine for Medical Diagnosis(2017)	QiangLi,Hailing Chen, HuiHuang, XuehuaZhao,ZhenNaoCai,	a new predictive framework is proposed by integrating an improved grey wolf optimization (IGWO) and kernel extreme learning machine (KELM), termed as IGWO-KELM, for medical diagnosis	Gains less number of selected features,achieving high classification performance	Grey Wolf optimization

Table 1: *Current Papers on Feature Optimization Techniques*

Wei Li et al, (2020) analyzed the influence of contour wave transformation on gray correlation degree and noise intensity of different medical images and improves the Bayesian threshold. The middle threshold function was improved by correlation characteristics of contour wave coefficients, and contour features of medical images were constrained by multiple thresholds. Based on the above, the dimension of the medical image was reduced by the wavelet multi-resolution analysis method, and the corresponding threshold search space was obtained. A ***genetic algorithm*** was used to find the best quasi threshold in the search space. Through this value, the attribute histogram of the medical image was established, the best feature extraction threshold of the medical image was obtained by the golden section method, and contour feature information of the medical image was extracted.

5. Chaos theory

Chaos theory (Devaney 1989) is the concept that a small change now can result in a very large change later. [18].It is said to be chaotic whenever its evolution depends on the initial conditions.Chaos helps the controlling parameter to find the global optimal solution more quickly and thus refine the convergence rate of the algorithm.

5.1Characteristics:

1.Extreme sensitivity to initial condition
2.Random like method found in non linear
3.Chaos is completely deterministic

5.2 Components:

- The time evolution of the equation
- Parameters that describe the system
- Initial condition

5.3 Chaotic Systems

Chaos is an unbalanced active behaviour that includes infinite unbalanced periodic waves in nonlinear ways, as well as significant requirements on the basic conditions. It has a logical core structure and three key active characteristics.

Initially, Ergodic property is the great stand that can search total nodes in the search plane of the system using formulas in definite range [7]. To avoid trapped problem in local optima, Chaos strategy is proposed to improve the value of searching global optimum in many optimization applications. So, the searching range [0, 1] can be consumed for feature selection system for optimizing the features by using chaos theory. This map with n dimensions is an isolated time dynamical method, which can be communicated using the below mentioned equation(10):

$$cu_i^{(k+1)} = f(cu_i^{(k)}),\ i=1,2,3\ldots n \qquad (10)$$

A chaotic order can be assessed using running the system function by defining the early state of $cu_i^{(0)}$,here chaotic orders can be described in the form of $cu_i^{(k)}$, k = 0, 1, 2, .The search function of chaotic evolution system is developed through calculating a chaotic vector value, which is decided using chaotic factor from a chaotic system that has the Ergodic property. The proposed system can develop its variants in many ways by applying a many of chaotic functions. This analysis present three kind of chaotic map systems such as logistic map, tent map and Gaussian map to examine the feature selection/optimization performances .

5.3.1 Chaotic Maps for Feature Selection

Chaotic method is considered to advance the performance of optimization algorithm to avoid trapped problem at the local optima also it increases the convergence speed. It is also employed to handle the random factors values of optimization algorithm . These maps show complex and dynamic behaviors, which occur in nonlinear systems and in determining system states .In this article , three Various chaotic maps are applied to increase the performance of feature selection which are described below.

5.3.2 Logistic map

This kind of map is a polynomial mapping with 2 degrees, which can rise from a chaotic occurrence by using a basic nonlinear system. The following equation demonstrates the mathematical expression,Eq(11)

$$X_n=\mu*x_{n-1}{}^{*(1-xn-1)}$$

Here μ is a factor, which adopts the performance of system; μ is typically set at (0, 4). This map does not venture interval [0, 1] into itself for the values of μ > 4, if μ = 4, the map system rises from chaotic occurrence. The split graph of logistic map is displayed in following figure 3, which offerings the entire system performance:

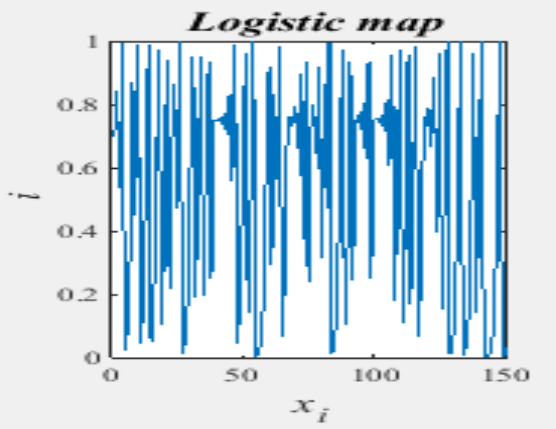

Figure 3. *Logistic Map View*

5.3.3 .Tent map

In this system, the parameter μ is classified (0, 2). The tent map proves a range of dynamical performances from probable to chaotic according to the value of μ. The map becomes chaotic when μ= 2,-Eq(12)

$$Xn = \begin{cases} \mu * x_{n-1} & \text{if } 0 \le x_n < 0.5, \\ \mu*(1-x_{n-1}) & \text{if } 0.5 \le x_n < 1 \end{cases}$$

This map is described as in above equation in which function curve is look similar to the tent form illustrated in below bifurcation figure .4,

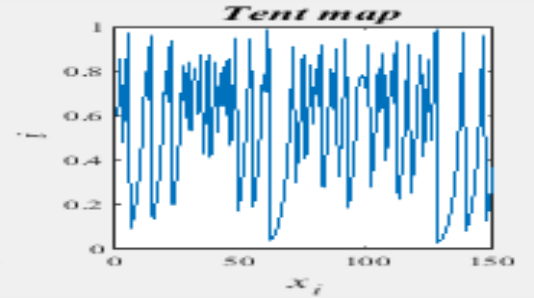

Figure 4. *Tent Map View*

5.3.4. Gaussian map

This is a nonlinear map, which is specified by using a Gaussian function operation equation(13)

$$X_n=\exp(-\alpha*x^2_{n-1})+\beta \qquad (13)$$

Gaussian map split graph look like a mouse presented in below figure 5. The map system can convert chaotic, once the parameters α and β are set at 6.20 and −0.5, correspondingly:

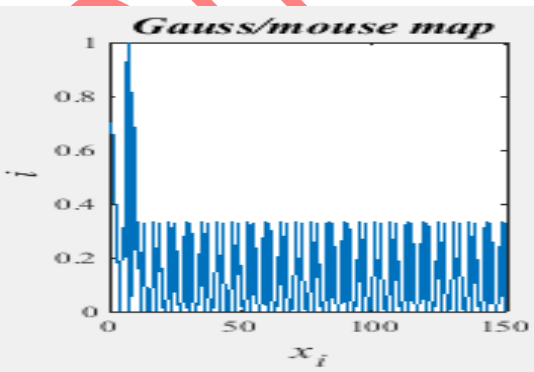

Figure 5.*Gaussian Map View*

5.3.5 Applications:

1. Robotics
2. Biology
3. Cryptography
4. Optimization Algorithms

6. Conclusion:

All the optimization algorithms are discussed in detail. These include: Genetic Algorithm (GA), Particle swarm optimization, Ant colony optimization ,Grey wolf optimization, and whale optimization

algorithm. Finally, chaotic systems are discussed which are extremely sensitive to their beginning conditions, meaning that even little changes in the starting point can result in massively different consequences. This makes the system fairly unpredictable. Future work will focus to apply chaotic version in different real time applications.

References

[1]. Youness Khourdifi1,Mohamed Bahaj1"Heart Disease Prediction and Classification Using Machine Learning Algorithms Optimized by Particle Swarm Optimization and Ant Colony Optimization" , , International Journal of Intelligent Engineering and Systems, Vol.12, No.1, (2019).

[2]. Jhila Nasiri and Farzin Modarres Khiyabani ," A whale optimization algorithm (WOA) approach for clustering ",Cogent Mathematics & Statistics, APPLIED & Interdisciplinary Mathematics (2018).
[3]. Majdi Mafarja, Iyad Jaber, Sobhi Ahmed & Thaer Thaher , "Whale Optimization Algorithm for high dimensional small-instance feature selection",International journal of parallel, emergent and distributed system(2019).

[4].Quoc-VietPham member ,IEEE ,SeyedaliMirjalili ,SeniorMember,IEEE,NeerajKumar, Whale Optimization Algorithm With Applications to Resource Allocation in Wireless Networks ,IEEE xplore Volume: 69, Issue: 4, April (2020).

[5]. B Sankara Babu1, A Suneetha, "Medical Disease Prediction using Grey Wolf optimization and Auto Encoder based Recurrent Neural Network" (2018).

[6]. QiangLi,HuilingChen,HuiHuang, "An Enhanced Grey Wolf Optimization Based Feature Selection Wrapped Kernel Extreme Learning Machine for Medical Diagnosis"(2017).

[7]. Zeynep Batik Garip,"The Chaos-Based Whale Optimization Algorithms Global Optimization", Chaos Teory And Applications Volume 01, Number 1,(2019).

[8].Yuhui Shi,"Particle swarm optimization:Development,applications and resources ",(2001),IEEE.

[9]. B Sankara Babu1, A Suneetha, G Charles Babu,Y , "Medical Disease Prediction using Grey Wolf optimization and Auto Encoder based Recurrent Neural Network (2018) ".
.
[10]. QiangLi, Huiling Chen,HuiHuang, XuehuaZhao, ZhenNaoCai, "An Enhanced Grey Wolf Optimization Based Feature Selection Wrapped Kernel Extreme Learning Machine for Medical Diagnosis(2017)".

[11]. Kanika Tyagi ,"A Comparative Analysis of Optimization Techniques, International Journal of Computer Applications" (0975 – 8887) Volume 131 – No.10, December2015.

[12].Longzhen Duan, " The optimization of Feature Selection Based on Chaos Clustering Strategy and Niche Particle Swarm Optimization" Hindawi Mathematical Problems in Engineering,2020.

[13].Arshad et al. "Chaos Theory and its Application: An Essential Framework for Image Encryption",2020.

[14].Y. Li et al. "Modified Whale Optimization Algorithm Based on Tent Chaotic Mapping and Its Application in Structural Optimization",Springer 2020.

[15].Yintong Li . "Whale Optimization Algorithm with Chaos Strategy and Weight Factor", Journal of Physics: Conference Series,2019.

[16].M.Kohli,S.Arora, "Chaotic grey wolf optimization algorithm for constrained optimization problems 2018.

[17].Fei Ye , " An improved chaotic fruit fly optimization based on a mutation strategy for simultaneous feature selection and parameter optimization for SVM and its applications",PLos One 2017.

[18]s. Biswas et al “Chaos Theory And Its Applications In Our Real Life”, Barishal University Journal(2018).

[19]. Asmaa Ahmed Awad1 , A “ Feature selection method based on chaotic maps and butterfly optimization algorithm”(2020).

Authors Biography

Dr.S.N.Geethalakshmi has completed MCA, M.Phil and Ph.D in Computer Science . She is working as Professor in the School of Physical Sciences & Computaional Sciences .Avinashilingam University ,Coimbatore. Her fields of research interest are Image Processing, Data mining and Software Project Management .She has published papers in the international journals and presented research papers in international and national conferences.

Ms.M.Janaki has completed MCA and M.Phil in Computer Science. she is currently pursuing her Ph.D in Computer Science in Avinashilingam University, Coimbatore. Her fields of interest are Data Analytics, Machine learning and Deep learning.She cleared UGC NET and SLET.