Archives of Computational Methods in Engineering (2023) 30:4831–4853
https://doi.org/10.1007/s11831-023-09963-4

REVIEW ARTICLE

A Comprehensive Review on RSM-Coupled Optimization Techniques and Its Applications

Anto Susaimanickam[1] · Premalatha Manickam[1] · Amalanathan Arputhasamy Joseph[2]

Received: 3 March 2023 / Accepted: 8 June 2023 / Published online: 23 June 2023
© The Author(s) under exclusive licence to International Center for Numerical Methods in Engineering (CIMNE) 2023

Abstract
This review article provides a comprehensive analysis of the optimization techniques used in a wide range of engineering applications. The comparison of various approaches such as Response surface methodology (RSM), Genetic algorithm (GA) and Artificial neural network (ANN) towards optimization problems is widely elaborated. The factors that affect the optimization using various techniques are addressed along with the safety precautions to be followed in a sequential manner to achieve a better optimization model. Furthermore, the coupling of two distinct algorithms (RSM-GA, ANN-GA) are explained and this hybrid approach provides a better localizing of the optimal point with a higher accuracy.

1 Introduction

Optimization commonly referred as a mathematical technique with a combination of scientific ideas and strategies for resolving a quantitative problem in various engineering fields [1, 2]. The term "Optimization" emerged as a result of the finding that quantitative issues in different engineering domains share a significant mathematical foundation with wide features occurring in common. Due to this similarity, the principle of optimization and its techniques can be used for designing and solving a wide range of problems. There are various steps to be understood by the researchers before addressing the optimization problem. It involves determining the requirement for optimization, allocating design variables, developing constraints and goal functions, establishing variable boundaries, selecting an appropriate optimization technique, and ultimately attaining the desired solution [3]. The process of determining the input variables that produce a function's maximum or minimum output is referred to as optimization. The continuous function optimization is one of the most widely occurring optimization problems in machine learning where both the input and output arguments are numerical values [4]. There are different algorithms available for optimization [5] which are categorized based on the informational data provided regarding the target function and towards the usage of gradient function. At the early stage, the optimization problems were solved based on the gradient information such as Bisection method, Gradient Descent and Newton's method [6]. Some of these algorithms can be used only for single input variable whereas others can be used for more than one input variable but with the existence of only a single global optimum. Generally, the gradient is obtained as a first step from the objective function which then performs the search of optimum value based on the step size. The step size used in the optimization algorithms affect the speed and accuracy of the results. A lower step size requires many data points but performs with a slower running rate and a higher step size runs more rapidly with fewer data points. Further, a lower step size results in minimal error on detecting the optimal point compared to higher step size. Thus, a lower step size is usually preferred in order to avoid the zigzag movement of the search space, thereby missing the probability of identifying the optimal point [7]. The major disadvantage with lower step size is a higher computation time which can be avoided by reducing the search space involved in the optimization algorithm.

Later, for complex objective functions where it was too difficult to find the derivatives, methods that were not reliant on gradient information (direct algorithms, stochastic, and population algorithms) were utilized [8]. The existence and accessibility of fast computing software are also one of the major reasons for the wide growth in different algorithms introduced towards the optimization problems [9]. In 1951,

✉ Amalanathan Arputhasamy Joseph
amalanathan.aj@hszg.de

[1] Department of Energy and Environment, National Institute of Technology, Tiruchirappalli, Tamil Nadu 620015, India

[2] Department of High Voltage Engineering, University of Applied Sciences, 02763 Zittau, Germany

Box and Wilson introduced the technique named Response Surface Methodology (RSM) which examines the relation between various input parameters with its associated output response [10]. The fundamental objective of RSM was to find the best response through a series of design experiments which are planned sequentially. The developers adopted a second-degree polynomial to perform this task although they were aware about the rough estimate. Still, they tried to explore as it was so simple for its evaluation and usage with less fact available about the mechanism. Later, researchers were much involved in developing a higher-level technique to provide an optimal solution that could solve a problem faster than the traditional algorithms and ended up in generating metaheuristic algorithms.

Metaheuristic algorithms in general, direct the search process by examining the search space to find the global optimum in lieu of local optimum values [11–13]. On the basis of search process criteria, metaheuristic algorithm is classified into (i) Metaphor based (or) Population based and (ii) Non-metaphor based (or) Neighborhood based algorithm. Succinctly, population based algorithm make use of multiple solutions in the process of searching whereas neighborhood based algorithm make use of single solution by means of local search [14, 15]. Variable neighborhood search (VNS), Tabu search (TS), Microcanonical annealing (MA), Guided local search (GLS) etc. are the popular neighborhood-based algorithms. On the other hand, Population based algorithms are nature inspired algorithms that could handle high dimensional optimization problems [16, 17]. This is sub-categorized into (i) Evolutionary computation, (ii) Swarm intelligence, (iii) Physics inspired algorithm, and (iv) Human inspired optimization algorithm. Evolutionary algorithms adopt the laws of natural evolution to provide a global optimum with significant and unbiased results. In this, population is initially created, and the algorithm parameters such as reproduction, crossover, mutation, and survivor selection are applied to obtain the optimal solution until it meets the termination criteria. Genetic Algorithm (GA), Genetic programming (GP), Differential evolution (DE), Evolutionary programming (EP), Evolutionary statistics (ES) are some of the notable evolutionary algorithms [18, 19]. Swarm Intelligence adopt the natural aspects of birds and mammals. Particle swarm optimization (PSO), Artificial bee colony (ABC), Honey bee mating optimization (HBMA), Ant colony optimization (ACO), Firefly algorithm (FA), Glow-worm algorithm (GWA), Dolphin optimization algorithm (DOA), Bat algorithm (BA), Cuckoo search (CS), Shuffled frog leaping algorithm (SFLA), Lion based algorithm, Monkey based algorithm, Wolf based algorithm, are the widely accepted swarm intelligence algorithms [20, 21]. All the above-mentioned swarm intelligence-based algorithms work on the natural behavior of each organism taken into account. It has its pros in finding faster optimal solution but delivers the local optimum values than converging to a global optimum because of its search process in a small space rather than in a large space [22]. Physics based algorithms adopt the laws of nature such as (i) Newton's law of gravitation, (ii) Quantum mechanics, (iii) Theory of Universe—Big Bang theory of expansion and Big Crunch theory where all matters are pulled by black hole, (iv) Electromagnetic systems, (v) Electrostatic systems—Coulomb's law, Gauss law, Newtonian law of Mechanics, Superposition principle of Electrostatics, (vi) Glass demagnetization, (vii) Galaxies. Simulated Annealing (SA), Gravitational search algorithm (GSA), Galaxy based search algorithm (GBSA), Charged system search (CSS), Atom search optimization (ASO), Sine Cosine algorithm (SCA), Henry gas solubility optimization (HGSO), Equilibrium optimizer (EO) [23, 24]. Strategy based on human's problem-solving intelligence, potential to comprehend, rationale, acquiring knowledge, ability to grasp and withhold ideas, supervisory and managerial powers etc. were set as basic inputs to build Human inspired optimization algorithms (HIOA). Many algorithms have been developed recently by acquiring the latest trends in human society. HIOAs such as Corona virus herd immunity optimization (CHIO), League championship algorithm (LCA), Harmony search (HS), Forensic based investigation optimization (FBIO), Political optimizer (PO), Teaching learning based optimization (TLBO), Heap based optimizer (HBO), Battle royale optimization (BRO), Human urbanization algorithm (HUA) are currently in research for solving optimization problem [22, 25–28].

After a span of 9 years since the introduction of RSM, Genetic algorithm (GA) gained importance with major research from Holland in 1960's [29]. As the name suggests, this algorithm was based on the natural selection theory from Charles Darwin. This algorithm was the first technique to adopt the system using multiple operators (crossover, recombination, mutation, and selection). Each of these operators forms a crucial role of the genetic algorithm model in terms of problem resolution. Compared to conventional algorithms, the genetic algorithm can handle complex objective functions parallelly. Since then, it has become popular not only in the field of biotechnology but also in the various other engineering fields (electrical, mechanical, aeronautical. etc.) where the optimization problems were more common [30, 31]. Along with the RSM and GA techniques, the artificial neural network (ANN) was also in its field of research in early nineteenth century [32]. Initially, only neurons with a single layer were used for its research until 1980. It was the work of Werbos on backpropagation which improved the neural network towards optimization problems [33]. At the beginning, only a specific activation functions were used for training the neural networks but then after the boom of deep learning technologies in 2010, there are multiple training algorithms and activation functions which efficiently modeled the system

[34]. There are certain drawbacks associated with these individual optimization techniques. The RSM has its restrictions with its user defined boundary conditions [35] while the GA involves multiple operators with higher computational time for its optimization. The training of ANN for its optimal model with different algorithms still requires more inspection for its better validation of response [36]. The above-mentioned algorithms are widely used for optimization problems in different fields separately and now the concept of coupling two different algorithms (RSM-GA, ANN-GA) are also tested for its performances in improving the accuracy. The coupling of two methods for optimization problems can deal with the problems exhibiting several local minima in the fitness function with its quick convergence towards the optimal solution [37, 38]. The main objective of present paper is to understand the principles of different algorithms towards the optimization problems, the parameters involved, logic of identifying the optimum point within its search space with a comparison of results from various domains of engineering.

2 Discussions

2.1 Response Surface Methodology

Response surface methodology, in accordance with the design of experiments implies the relationship between the response/output variables of interest and the associated/input variables through a set of mathematical and statistical techniques. Since its inception in the early 1950s, it has been at the forefront in research and industrial experimentation fields [39, 40]. Here, the response is the dependent variable and the parameters that affect the response are the independent variables. Optimization process through RSM is obtained through certain stages which is described in Fig. 1. Even though the relationship between them is concealed, it can be approached by a low degree polynomial model which is represented as

$$y = f\left(X_1, X_2, X_3 \ldots \ldots ..X_n\right)\beta + e \tag{1}$$

where y is the response and the function $X_1, X_2, X_3 \ldots \ldots ..X_n$ is the independent variable with β as coefficient and e is the experimental error [41]. From Eq. (1), determining the function f implements the prediction of response for any values of X that are not included in the experiment. The representation of $f(X_1, X_2, X_3 \ldots \ldots ..X_n)$ is called a response surface. The approximation of the response function is called RSM. However, if the function f is known, then the values of X can be obtained by the calculus method to give the optimum response. But, in most of the scenario, the mathematical form of function f is unknown. In these cases, the method of approximation is applied within the stipulated experimental region by the polynomial degree.

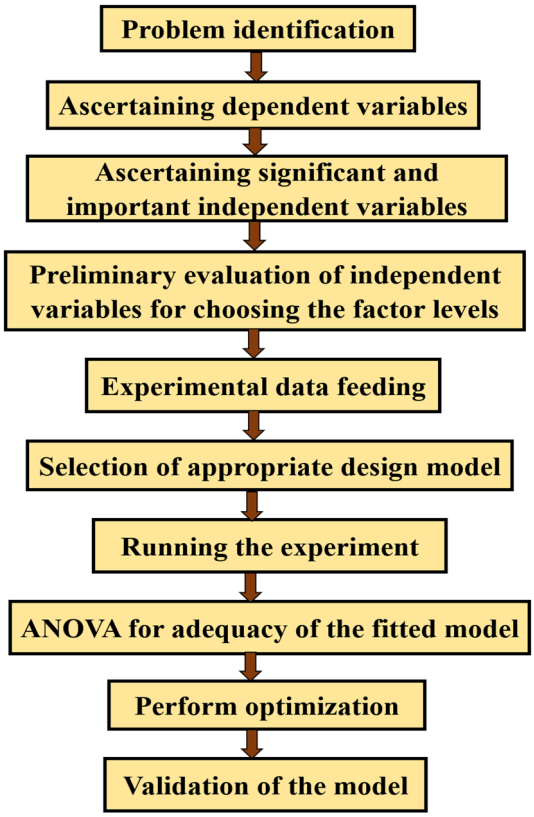


Fig. 1 Flow chart of RSM

This method of approximation happens when the independent variables are fed as inputs and the corresponding outputs (response) are estimated according to the specified function that exhibits between the response and the independent variables. It then analyses the values of inputs with respect to the responses and performs the approximation in order to determine the optimum response. The approximation model technique does this work. Hence, if there exists a linear relationship between response and input values then the method of approximation is a linear model, whereas if there are highly non-linear outputs with respect to input values, then it follows the cubic model approximation technique. Likewise, few model approximation techniques are available and will be executed based on the input and response values. The optimum response within the experimental region is found out by eliminating the low significant terms and by minimizing the fitting errors that occurs during the application of approximation model technique. This can be done by sequential replacement, stepwise replacement, and exhaustive search methods [41, 42]. If the function takes the degree of 1, then it is known as first degree model. In this model, the response obtained is linear fashion with the independent variables and is given by Eq. (2) [43, 44]

$$y = \beta_0 + \sum_{i=1}^{n} \beta_i + e \tag{2}$$

On the other hand, the function with two degrees is known as second degree model. In this second degree model, the response obtained is in the form of curvature which is represented as Eq. (3) [45, 46].

$$y = \beta_0 + \sum_{i=1}^{n} \beta_i + \sum_{i=1}^{n} \beta_{ii} x_i^2 + \sum \sum_{i<j=2}^{n} \beta_{ij} x_i x_j + e \tag{3}$$

2.1.1 Experimental Designs for Fitting the Model

From a set of unorganized data, getting the accurate fitting of response designs is quite computationally complex and involves the precision approach for estimating the response. In such cases, designs for fitting the model could be a better option wherein the independent variables of any combinations can be fed as input for generating the data in estimating the response [47, 48]. The designs for first- and second-degree models helps in analyzing the correctness of the response and the region of appropriate response within the experimental region. The first degree designs available in the literature are (i) 2^k factorial design, (ii) Plackett Burman design, (iii) Simplex design [49, 50]. The second degree designs include (i) 3^k factorial designs, (ii) Central composite designs, (iii) Box-behnken design [51, 52] (Table 1). Second degree model takes the general form of Eq. (3) as mentioned in the manuscript. It is noticeable that the second-degree model contains (i) linear terms, (ii) pure quadratic terms and (iii) interaction terms. This second order model is used when there exists a response surface in the form of curvature. The response obtained using this second degree model can be in any one of the standard shapes (referring to surface 3D plot). For example, an upward curve indicates that the model has found the apparent maximum value, and a bowl shaped curve indicates the apparent minimum value and there also exists a minimax system where they exhibit both minimum and maximum behaviour. The logic behind this optimization is the analogy of Taylor series (used to approximate the complex functions). First order Taylor series is analogous to first order regression model, and second order Taylor series is analogous to second order regression model [53].

Response surface methodology (RSM) is a key optimization approach and it has undergone substantial progress to address the concerns in multiple engineering domains. Based on the observational data reviewed from various literatures [47], around 14.7% of research works in engineering have used RSM as its optimization tool. Among them, a few research works have been discussed in this Section. The saline wastewater treatment adopting electrochemical oxidation process was studied towards the removal of chemical oxygen demand (COD) and total organic carbon (TOC) efficiency [74]. The experiments were performed considering the parameters of pH, applied voltage, salt concentration and reaction time used for the process. The prediction of the model adopted was higher than 0.95 towards the output responses and it was inferred to a better tool for identifying the optimal parameters towards the wastewater treatment. The effectiveness of biodiesel as an alternative fuel for compression ignition engines has been researched for many years but still the commercialization of biodiesel as a fuel is constrained due to higher cost involved. Kusum oil has been utilized to make improved biodiesel [75] using the process of transesterification to address the aforementioned concerns. The yield in the biodiesel along with other physio-chemical properties were the responses and the experimental model was predicted better using RSM model and was found to be acceptable as per ASTM6751 standard listed for biodiesel. Coal is one of the most predominant energy resources in the world due to higher availability and its affordability towards the prices. Due to scarcity in the higher-grade coals, the lower grade coals must be used to meet the energy demands of the present load conditions. The lower graded coal has a significant impact on the process towards higher ash and moisture content, thus Behera et al. [76] investigated the process of reducing ash content from lower grade coal with variables such as temperature, time and acidic concentrations. The optimization was carried out using central composite design (CCD) methodology and it was inferred that acidic concentrations was more effective in minimizing the ash content compared to impact on time and temperature. Nanofluids are currently the most relevant field for researchers due to its widespread use in business as well as in technology for its improvement in heat transfer. Hatami [77] formulated a wall to understand the impact of nanofluids from improving heat transfer mechanism using RSM model to obtain an increased Nusselt number. The findings of the research concluded an optimized diameter upto 1.0 can enhance the natural heat convection with better Nusselt number. The depletion of petroleum-based resources has not only brought a change in the modes of transport from moving towards electric vehicles but also led to the development of alternative fluids for transformer applications. Any new fluids before being introduced into transformer should be tested for its better dielectric properties. Thus, the transesterification of Pongamia Pinnata oil (PPO) was performed considering reaction time, temperature and catalyst where its response towards the breakdown voltage, viscosity and fire point were modelled using RSM [78]. The quadratic equation was used for its responses which resulted in better significance on the analysis of variance (ANOVA) and could be used as

Table 1 Types of experimental designs in response surface methodology

Experimental designs for model fitting	No. of coded levels	No of experimental runs	Design property	Applications	References
First order degree designs					
2^k factorial design	Screening of variables/factors with two coded levels, namely − 1 and + 1 that corresponds to low and high level of factors respectively	2^k runs where k is the no. of factors. If k = 3, then no. of experimental runs is 8	Orthogonalilty (Property where the individual effects of each independent variables can be estimated without or with mere effects of a third variable that influences the effect of independent variable on dependent variable)	(i) Initial screening of variables (ii) Forms the basis of construction of second order response surface designs like CCD, BBD (iii) Useful in fitting the first order by the method of steepest ascent or descent	[54–56]
Plackett Burman design	A non-geometric designs that has two coded levels namely − 1(low level) and + 1(high level)	K = N − 1, where N is the no. of experimental runs. Therefore, if we have 10 variables, then the no. of experimental runs is 11. It gives best results when the runs are in multiples of 4 i.e., 4, 8, 12, 16, 20 etc.	In this method, there is no significant interactions between the variables, however with the introduction of hidden projection property, the interaction effects can be estimated. This projective property has two aspects (i) Geometric projection and (ii) Ability to estimate interactions. Therefore, a Plackett–Burman with their hidden projection property is a result of complex patterns between the interactions and the main terms	(i) Widely used for estimating main effects (ii) Screening of important factors by eliminating the least or unimportant factors in determining the response (iii) Effective for two level multi-factor experiments	[57–60]
Simplex design	A design of equally/uniformly spaced points on a lattice with three coded levels	K + 1, where k indicates the dimensions and k + 1 indicates the vertices. If K = 2, then it forms a equilateral triangle, and if K = 3, then it is a tetrahedron	Orthogonality	(i) Estimates every independent variables and their interactions (ii) Determines the best proportion of variables that give rise to desirable response	[61–63]
Second order degree designs					
3^k factorial design	It includes all the permutations of k variables in all levels. It has three coded levels namely 0 (low), 1 (intermediate), 2 (high)	3^k, where k is the no. of variables/factors. If k = 3, then the no. of experimental runs is 27	Orthogonality	(i) Helps in studying many factors/variables simultaneously (ii) Estimate both the main effects and the interaction effects that give rise to desirable response	[64–66]

Table 1 (continued)

Experimental designs for model fitting	No. of coded levels	No of experimental runs	Design property	Applications	References
Central Composite Design (CCD)	It is constructed based on 2^k factorial design which has five coded levels namely $-\alpha, -1, 0, +1, +\alpha$. A indicates the axial/star points, 1 indicates the factorial points, and 0 indicates the center points. A is calculated by $2^{k/4}$, where k is the no. of variables/factors	No. of experimental runs is indicated by $2^k + 2k + k_c$, where 2^k indicates the factorial points, 2 k indicates the axial/star points, and k_c indicates the center points. Therefore, if there are 3 factors, then the no. of experimental runs are 20 if the center point is repeated 6 times. However, the number of center points influences the design functions. The prominent reasons to perform experiments with default runs are, (i) The default number of center points provides near uniform precision designs, (ii) Replicated center points are used to estimate pure error for the lack of fit test [67]	Orthogonality and rotatability	(i) Helps in finding out the optimal solution (ii) Effective in studying main and interaction effects (iii) Quick, precise, robust identification of first or second order terms	[68–71]
Box Behnken Design (BBD)	It is the combination of 2^k and incomplete block design which has three coded levels namely, $-1, 0, +1$	$2k(k-1)+1$ is the formula for finding out the total experimental runs in BBD. If there are 3 factors/variables, then no. of experimental runs is 13. Here, the no. of experimental runs is less than CCD	It is designed to be either rotatable or almost rotatable	(i) The no. of experimental runs is less than CCD (ii) Helps in finding out the main and interaction effects	[63, 72, 73]

a suitable tool for optimizing the parameters involved in transesterification of insulating oils towards transformer applications.

2.2 Genetic Algorithm

Genetic algorithm is a stochastic search algorithm that adopts the principle of "survival of the fittest" and the elimination of unfitted individuals. Since GA is adopted from the principle of natural selection and reproduction from biological processes in nature, there are many terms in GA that are comparable with the biological terms. As defined by the natural biological process, optimization involved in GA are random. However, GA let us to set the level for randomization and have control over it [79–81]. The following terms can be understood from Fig. 2. (1) Population, (2) Chromosome, (3) Genes, (4) Allele.

Population is all possible solutions for the specified problem represented by the set of individual chromosomes, computationally represented as bit strings (assignment of binary numbers to the chromosome). In that, one possible solution (a bit) to the specified problem denotes chromosome, which is comprised of genes. The value given to each gene represents the allele [82]. In addition to that, population in actual (or) real system must be converted to the population that is easily understood in the computational space. Genotype and Phenotype are the two terminologies that represent the population in computing system and the population in actual system respectively. Encoding and Decoding are the mapping systems that transforms from phenotype to genotype and from genotype to phenotype in a design space respectively (Fig. 3).

The general structure of GA (Fig. 4) for solving optimization problems start with the random initialization of population and each of the chromosomes in population is evaluated by fitness function and follows the termination path to reach the optimal solution. If the termination criteria are not met by the chromosomes, then new population is generated by applying the GA operators such as selection, crossover, mutation [83]. It is then evaluated by fitness function and checked for its termination criteria to obtain the optimal solution. The process of generation of new population with the help of GA operators occur until the termination criteria is satisfied. This is the general framework of GA [82]. Naturally, GAs has a large number of parameters that must be modified to achieve the optimal performance for any optimization challenge.The important parameters that one should know for performing any optimization problems using GA are (i) Fitness function, (ii) selection, (iii) Crossover, and (iv) Mutation.

2.2.1 Fitness Function

It is one of the main input parameters that defines the optimization problem. In other words, it is the only information available for solving the problem [84]. For any kind of optimization, say for instance minimization or maximization problem, fitness function is needed to evaluate the

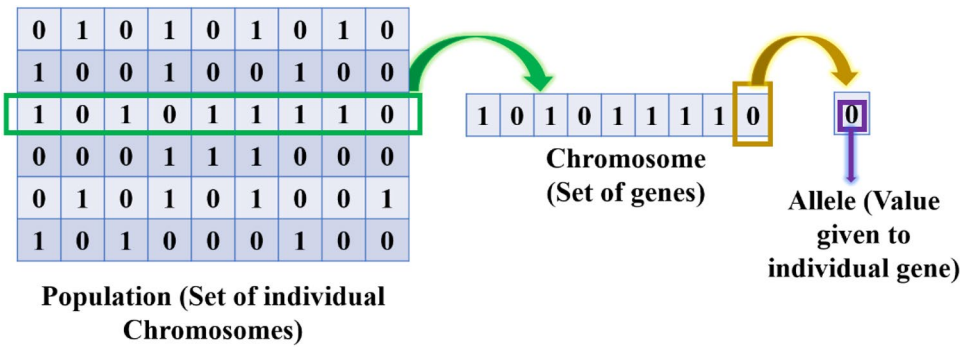


Fig. 2 Primary terminologies in Genetic Algorithm

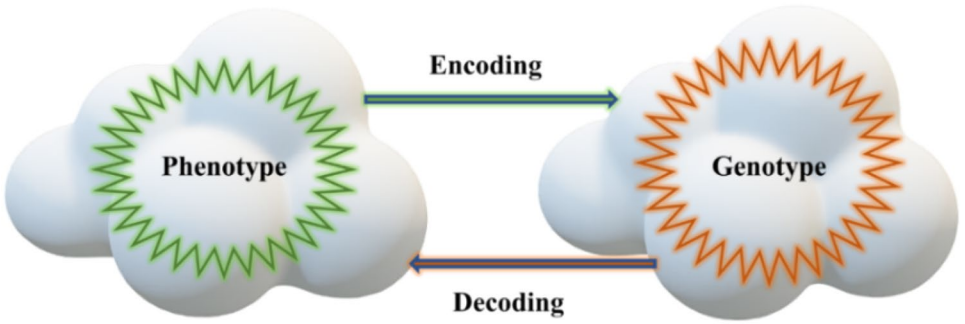


Fig. 3 Mapping systems in Genetic algorithm

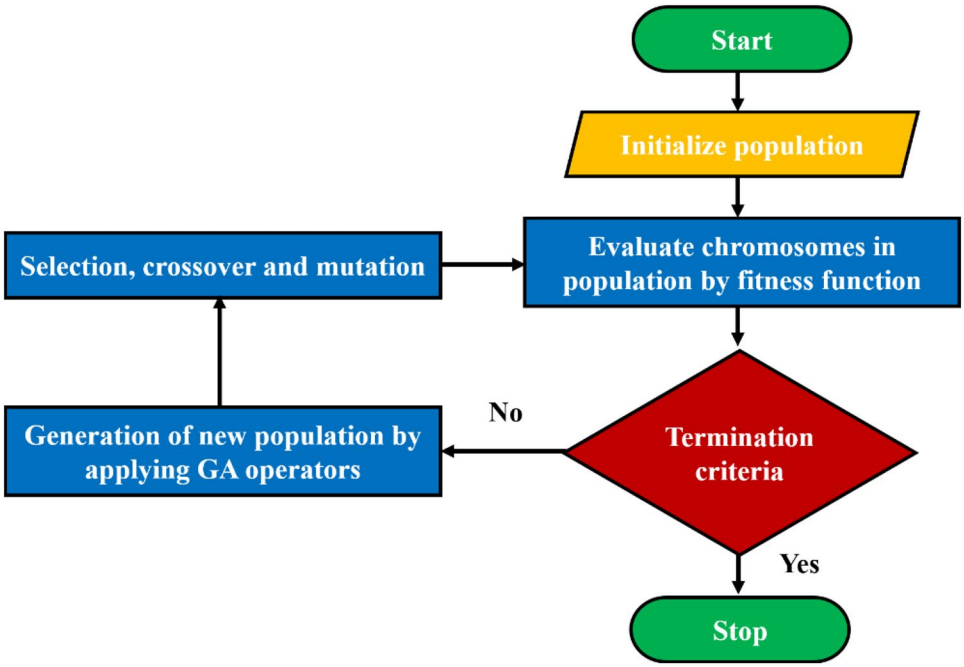


Fig. 4 Structure of Genetic Algorithm

individual chromosome in a population which ultimately gives a fitness score. This fitness score helps in finding out the best individual chromosome for giving optimal solutions or improvisation by successive iterations using GA operators [85, 86]. The fitness function for each problem is different. According to the stated problem, a particular fitness function should be applied. The trickiest element of creating a challenge for genetic algorithms is coming up with a fitness function for the given situation. Error measures like Euclidean distance and Manhattan distance are frequently employed as the fitness function for classification tasks involving supervised learning. Whereas basic functions can be employed as the fitness function for optimisation problems, such as the sum of a group of computed parameters relevant to the problem domain. In simple terms, fitness function can be explained as shown in Eq. (4).

$$\text{Function } y = f(x) \tag{4}$$

For instance, $y = 2\left(x^2 - 3x\right)$where fitness function is $2\left(x^2 - 3x\right)$. This indicates that the output variable y is dependent on x variable and in order to know at what value of x, the function y is minimized or maximized, optimization is performed and the plausible way of solving is through GA. The value of x is represented as the solution which can be either identified as the solution for maximization or minimization problems. For any optimization problems, fitness function takes the value of each parameter that is known to contribute the output i.e., it can be a sum of all individual parameters that contribute towards a positive output or it can take parameters that have positive and negative effects on the output. In the fitness equation, square terms take the positive coefficient whereas the individual term takes a negative coefficient. In optimization problems, objective value is same as fitness function that is of great importance in determining the output or response, provided the factors affecting the response is specified in the equation [87]. Hence to get the global optimum value, each solution is evaluated through the fitness function. The best solution after fitness function evaluation is then subjected to genetic operators such as crossover and mutation [88]. Recently, many methods have been implemented in promoting the accuracy of fitness function evaluation in finding the global optimum. A, K-means index (KMI), Partition separation index (PSI), Separation index (SI), Davis–Bouldin index (DBI), fuzzy c-means index (FCMI) [89, 90], Gaussian process [91], Artificial neural network (ANN) [84].

2.2.1.1 Selection It is the essential parameter in choosing the best solution after fitness function evaluation. It is also referred as reproduction operator [92] because production of new individuals depend on the selection criteria and the possible solutions are obtained through reproduction and crossover [93]. The general strategy employed during the process of selection is that the individual with the highest fitness score is selected and copied for creating new population whereas the individual with least fitness score is eliminated [81]. But it is not assured that the highest fitness score always builds the global optimum solution. To the contrary, the least fitness score individual can also contribute reach-

ing towards the optimum. Having said that, an appropriate selection strategy must be employed such that the least fitness score individual is not completely eliminated and is taken into account for selection [94]. The selection operator is used to choose solutions from the current population to build the next population of solutions, which serves as the foundation for the algorithm's subsequent iteration. However, variety from crossing and mutation must be balanced with selection. Strong selection will result in the dominance of highly fit individuals in the population, diminishing the diversity necessary for innovation and advancement. On the other hand, extremely weak selection may cause overly sluggish evolution [53]. Selection methods that are available in literatures are (i) Roulette wheel selection, (ii) Stochastic universal sampling, (iii) Linear rank selection, (iv)

Table 2 Selection techniques/methods in Genetic algorithm

Selection method	Description	References
Roulette wheel selection	(i) Also known as fitness proportional selection (ii) In this selection method, the wheel containing different individual is spun and a fixed pointer is placed for selection (iii) The wheel is partitioned according to the size of the individual i.e., larger the size of individual, larger area it takes in the wheel and vice versa. So, when the wheel is spun selecting the larger area is of higher probability (iv) The process of selecting individual proceeds till the smallest individual is pointed (v) Therefore, the probability of individual selection is directly proportional to the fitness function score and is given by $P_i = \frac{f_i}{\sum_{j=1}^{n} f_j}$ (5) where f_i is the fitness score of ith individual and N is the size of population	[93, 95]
Stochastic universal sampling	(i) Variant of Roulette wheel selection which was developed to overcome the convergence of solution at local optima in roulette wheel selection (ii) In this, N equally spaced pointers are employed for selection where N is the number of individuals to be selected and the space between the pointers is given by 1/N (iii) However, the individual is partitioned according to the size like in the case of roulette wheel selection (iv) Hence, random individual is selected based on the position of pointer	[96, 97]
Linear rank selection	(i) In this method, each individual is assigned a rank based on the fitness score (ii) Consider a population of size N, rank N is given to the best individual followed by N-1, N-2, …1 to the least individual (iii) This method also overcome the convergence issue to local optima (iv) The probability of selection is given by $P_i = \frac{Rank_i}{N(N-1)}$ (6)	[94, 98]
Exponential rank selection	(i) his method follows the same procedure of linear rank selection for assigning ranks to the individual (ii) However it differs in the selection probability criteria (iii) Here, probability is exponentially expressed and is given by $P_i = \frac{C^{N-i}}{\sum_{j=1}^{N} C^{N-j}}$ (7) where C is the base of exponent; $0 < C < 1$ (iv) Compared to linear rank selection, the chance of eliminating the least individual is considerably less, thereby leading to select global optima	[94, 95, 99]
Tournament selection	(i) Most prominent selection technique in genetic algorithm (ii) In this technique, N individuals in a population are selected at random and are ranked based on the fitness score (iii) The number of N individuals are referred as tournament size (iv) The highest fitness score of individuals from the tournament size gets selected for further process (v) Hence, larger the tournament size, higher is the chance of selecting the highest fitness score individual, thereby leading to eliminate the lowest fitness score individual (vi) In most cases, tournament size is kept as 2 (Binary tournament selection) to avoid the convergence to local optima	[94, 99]
Truncation selection	(i) One of the simplest techniques that is based on the fitness scores of the individuals (ii) Truncation threshold is the main factor in selecting the individuals which ranges from 50%-10% (iii) This percentage denotes the proportion of population to be selected and any individual below this threshold is eliminated (iv) This often results in stuck with local optima	[100–102]

Exponential rank selection, (v) Tournament selection, (vi) Truncation selection [93, 95] (Table 2).

2.2.1.2 Crossover After the selection of best individuals, crossover genetic operator is applied in order to create a new individual that derives the properties of best individuals (Umbarkar and Sheth). In simple terms, production of new offspring from the best parents by means of exchanging the genes is referred as crossover. This finds an advantage in finding out the new individual whose characteristics are way better than their parents. There are many classifications of crossover operators available in the literature which is represented in Fig. 5 [103–108].

2.2.1.3 Mutation Mutation is an evolution operator that helps in modifying the genes in a chromosome after the crossover operator is applied [109]. This holds true because at times, the new individual produced from crossing over may stuck at local optima, even if their parents are the best individual. Mutation operator finds advantage in genetic algorithm by exploring the search space when the crossover operator makes use of it to find the best individual. Here, the individual is modified based on the mutation probability [110]. By enabling mutation operator, it is known that the diversity of the entire population is maintained and avoids the convergence to local optima [111]. In literature, there are different types of mutation operators but is not limited to the mutation operators in Table 3.

The genetic algorithm has undergone extensive research, testing, and use in numerous engineering domains. It not only offers an alternate approach to problem-solving, but it also outperforms the other conventional approaches in the majority of the issues studied [122]. Many of the real-world issues of determining the best parameters could be challenging for conventional approaches, but they are perfect for genetic algorithms [123]. Over the past few decades, many intrusive modeling techniques have been formulated for granular soils. This has led to a significant increase in the difficulty of choosing an acceptable model with the required features based on standard testing and with an efficient method for determining the factors involved in geotechnical fields. So, in order to examine an appropriate sand model, Jin et al. [124] investigated the suitable parametric detection of both drained and undrained testing process. The estimation of minimal objective parameters along with lower strain levels were determined based on the GA optimization where the experimental and simulation results provided a

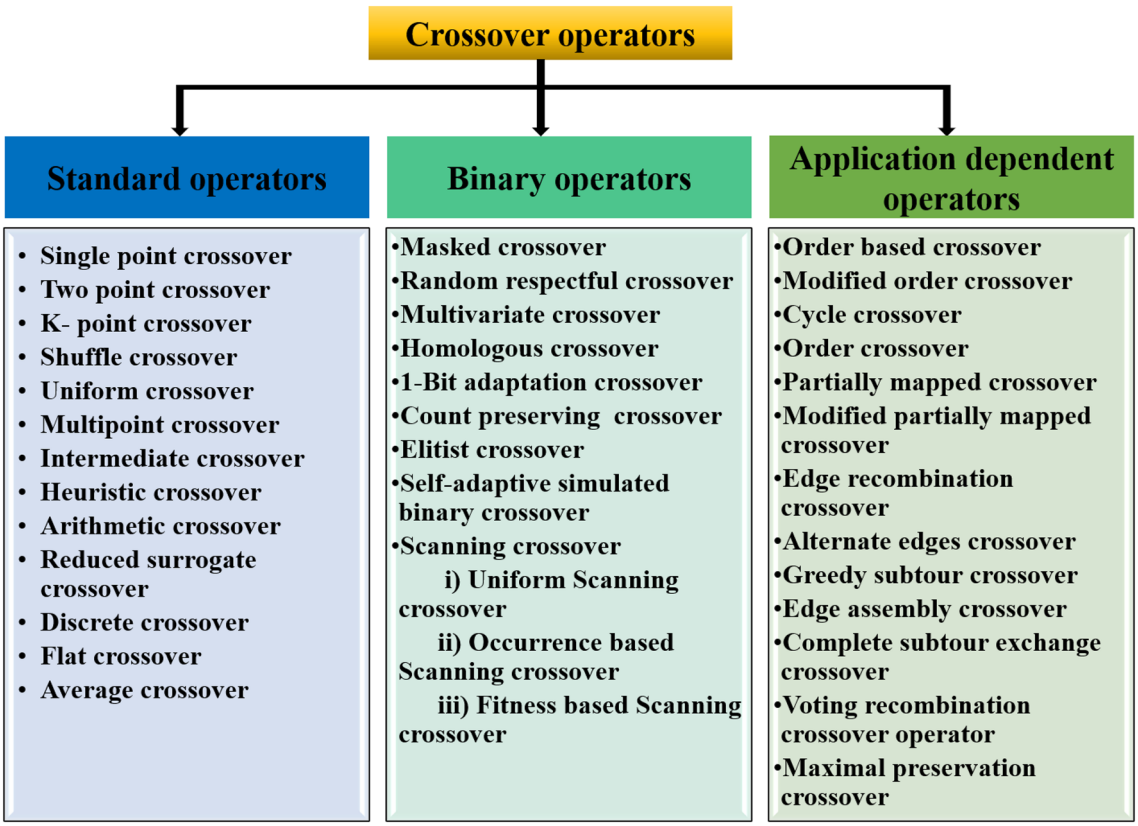
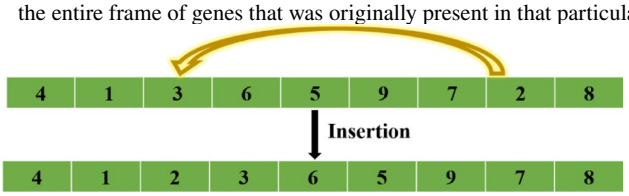
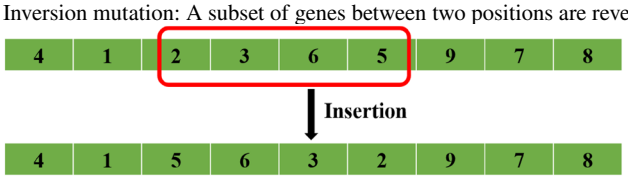
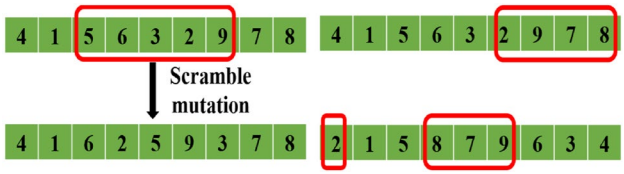
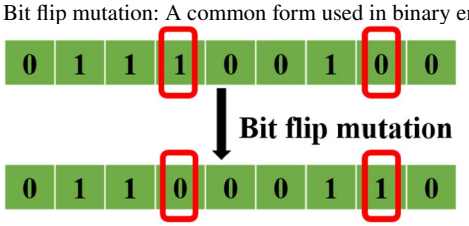
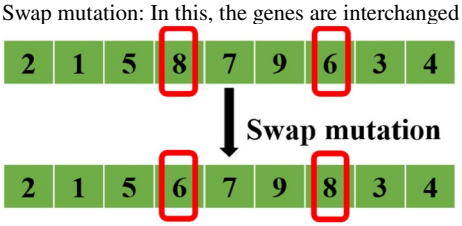
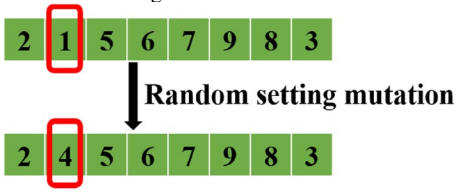


Fig. 5 Classification of crossover operators

Table 3 Types of mutation operators in genetic algorithm

S. No.	Mutation operators	References
1	Insertion mutation: In this, a gene in a chromosome is selected and randomly inserted into a new location, thereby shifting the entire frame of genes that was originally present in that particular location 	[112, 113]
2	Inversion mutation: A subset of genes between two positions are reversed and placed in that position 	[114, 115]
3	Scramble mutation: In this, subset of genes is selected, scrambled and placed either in the same position or in a new position 	[113]
4	Bit flip mutation: A common form used in binary encoding. The bits 0 and 1 are replaced in the chromosome by 1 and 0 	[116, 117]
5	Swap mutation: In this, the genes are interchanged with each other 	[118, 119]
6	Random resetting mutation: Random values from the given range is applied to the existing genes 	[120, 121]

good correlation. Reverse osmosis (RO) has a broad range of industrial applications as a separation technique in comparison with traditional thermal processes [125]. The pilot-scale model for RO have been tested for its removal of chlorophenol from the waste-water [126] and the optimization using GA was created with the objective of increasing the chlorophenol rejection and minimizing the operating pressure conditions. The findings indicated a chlorophenol rejection of upto 26.57% with its pressure maintain within its limits. Algal biofuels are gaining popularity in the effort of cutting down the carbon emissions in the environment, but it is still unclear on how the fuel generations are sensitive to different variables [127]. Azari et al. [128] studied the impact of different parameters (aeration, time, light intensity, pH) towards the rate of CO_2 biofixation of Chlorella vulgaris. The prediction of experimental dataset achieved using GA was around 93% and the same model is to be understood for large scale experimented to know the limitation of the techniques towards the commercial scale of biofuels. Introducing the renewable energy generations (solar, wind) into the grid using power electronic devices needs a perfect synchronization and thus the smart grids are developed which provide a better metering of load at the distribution level and indicating the level of harmonics at the transmission levels [129]. The integration of different generations is to be allocated with economic loading and Arabali et al. [130] have used the integration of solar and wind generations along with energy storage devices on meeting the HVAC loading. The suggested methodology using GA optimization is a perfect tool for energy management which can be used by the utility companies to combine the different energy generations on meeting the various loads such as residential, commercial, and industrial feeders with an optimal cost.

2.3 Artificial Neural Network

The Artificial neural network (ANN) performs the functional relation between the input and output similar to the biological neural system in the human body [131]. The interconnection of different neurons is processed through wide range of layers. The typical architecture of ANN model with single hidden layer is shown in Fig. 6. The ANN involves three different layers: first layer is the input layer where the experimental data points are provided to the network, the second layer is known as hidden layer where the input neurons are allowed to perform transformations using an activation function and the third layer is the output layer which calculates the responses from the network. Depending on the regression performance, the number of hidden layers can be increased to improve the accuracy. With increase in the hidden layer, the complexity of the model also increases. So, generally the researchers test the network with an optimum level of hidden layers [132] to reduce the difficulty involved

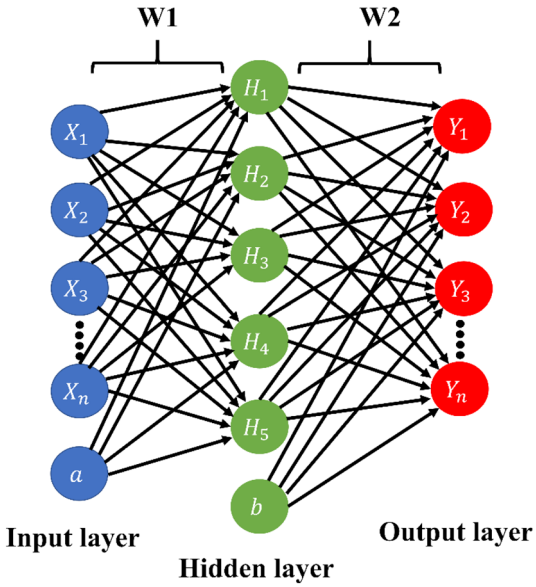


Fig. 6 Typical Architecture of ANN model

in the formulating the objective function and at the same time maintain a higher precision. The ANN performs training, validation and testing process from the input and output data informations. Each of these steps could be assigned with certain percentage from informational database to perform the neural network. Mostly, the training phase requires a higher information compared to validation and testing phases [133, 134]. From the neural network formed, the Eq. (8) could be defined as:

$$Y = \sum_{j=1}^{n_h} \left\{ f_2 \left[W2_{j,1} * \left\{ \sum_{i=1}^{n_i} \sum_{j=1}^{n_h} f_1 \left(X_i * W1_{i,j} + a_j \right) \right\} \right] + b \right\} \tag{8}$$

where n_i is the number of neurons in the input layer, n_h is the number of neurons in the hidden layer, f_1 and f_2 are the functions used between input layer to hidden layer and hidden layer to output layer, $W1$ and $W2$ are the weights between input layer to hidden layer and hidden layer to output layer, a and b are the bias added to the weights in the input and hidden layers, X indicates the input variables and Y is the output responses respectively.

The human brain works for each second on categorizing the informations as useful and non-useful groups. Similar phenomenon is exhibited in ANN where the segmentation process helps in utilizing the required information with the activation function module [135]. The main role of the activation function is to convert the input variables which are weighted along with a bias towards the hidden layer or output layer. The different hidden layers in the neural network will utilize the similar activation function while the output layer using a different activation depending on the

prediction of the neural network model [136]. Without the use of activation function, the neural network would only perform a linear transformation from the weights and bias added to the input variables. Thus, the activation function will introduce a non-linearity in network during its feedforward propagation. There are different activation functions (Binary step function, linear activation function, non-linear activation functions) which are used in ANN [137, 138]. The binary step function compares the input with a threshold value to determine the activation process and the linear activation function is termed as no activation function since it only does linear transformation towards the input variables [139]. Both binary step function and linear activation function has a gradient as zero and constant value which has no relation towards the input variables, and hence could not be used for backpropagation [140]. Further, irrespective of the number of hidden layers used, the neural network is reduced to a single layer when the activation function is linear. The non-linear activation function overcomes the above limitations by allowing backpropagation due to its gradient information and makes it feasible to determine which weights present in the input neurons could be modified to provide a better prediction of the output response [141]. The different non-linear activation functions used in the neural network are sigmoid or logistic function, hyperbolic tangent function, Rectified Linear Unit and Exponential Linear Unit [34, 141]. Among the above-mentioned functions, the sigmoid and hyperbolic tangent functions are widely used. The non-linear activation functions are used in the hidden layer whereas the linear activation functions are used in output layer. The ANN is a supervised machine learning techniques that uses different training algorithms such as Levenberg Marquart (LM), Bayesian regularization (BR) and Scaled Conjugate algorithm (SCG) for training the network [142, 143]. LM algorithm is the fastest training algorithm with higher requirement of memory space. BR algorithm minimises the combination of squared errors and weights to identify the perfect model for its generalization of network. SCG algorithm uses gradient techniques for its update on weights and bias which are more efficient towards large problems with lesser memory requirement than Jacobian calculations involved with LM and BR algorithms. Each of these algorithms formulates a unique methodology on framing the ANN model that affects the precision during training process [144] which is brief discussed below:

2.3.1 Levenberg Marquart Algorithm

The Levenberg–Marquardt algorithm was developed to operate on loss functions that takes the form of sum of squared errors and it functions without calculating the correct Hessian matrix. On the contrary, it utilizes the Jacobian matrix and gradient vector for its calculation. The approximation of Hessian matrix (H_{LM}) with a second order derivatives and gradients (g) is shown in Eqs. (9) and (10) once the parametric function takes the shape of sum of squares [145].

$$H_{LM} = J^T J + \lambda I \tag{9}$$

$$g = J^T e \tag{10}$$

where J is the Jacobian matrix which involves the first order derivative of error involved in the network regarding the bias and input weights, λ is the regularizing parameter, I is the identity matrix and e represents error obtained from the network. The parameter λ forms a major role in the functioning of LM algorithm. If λ is set to zero, then Eq. (9) of LM algorithm approaches the Newton method and if λ is assigned a larger value, the LM algorithm are effective as a gradient descent algorithm [146]. It is much simpler to calculate the Jacobian matrix compared to Hessian matrix using a conventional backpropagation method. The LM algorithm utilizes these approximations towards the Hessian matrix calculation in the Newtons method for updating the values at each iteration. The network performance towards error will be reduced during each iteration and this approach of LM algorithm accelerates the convergence and makes it very fast while training the neural network compared to traditional gradient methods [147].

2.3.2 Bayesian Regularization Algorithm

Bayesian regularization (BR) algorithm based on Bayes' theorem is a mathematical process that converts a non-linear regression into a statistical problem [148]. The benefits of Bayesian regularized artificial neural network (BRANN) is the predictions towards the model to be more robust eliminating the need for validation procedures. These networks provide solution to a wide number of problems involved in quantitative structure activity relationship models. They are challenging to overtrain because data techniques offer a Bayesian objective criterion for ceasing the training. They are especially difficult to overfit due to the BRANN's ability to calculate and train on a variety of useful network characteristics or weights, effectively turning off those that are not significant. The number of weights in a traditional neural network algorithm is typically much larger than BRANN [142]. The Bayesian regularization (BR) algorithm adjusts the input bias and weights in accordance with the optimization from LM algorithm [149]. A perfect generalization of the network is created by minimizing the weights and squared errors. The network weights are being introduced in the fitness function ($F(\omega)$) as shown below [150]:

$$F(\omega) = \alpha E_\omega + \beta E_D \tag{11}$$

where E_ω and E_D are the sum of squared network weights and errors with α and β representing the parameters of the objective function. Once the optimum values of these parameters have been identified, the algorithm is switched towards LM technique for its calculation on Hessian matrix and updated weights are used for the minimization of objective function.

2.3.3 Scaled Conjugate Algorithm

The conjugate gradient algorithms execute their searches in a way that generally leads to a convergence faster than the steepest descent method despite maintaining the error minimization attained in the previous iterations [151]. The calculation of step size is performed using line search method rather than the computation of Hessian matrix for determining the optimal distance to advance in the direction of search space. Further, the step size will be altered during each iteration such that it minimizes the objective function in the direction search of the conjugate gradient. In addition to line search method, there are various methods that can be used for estimating the step size. The concept was to integrate the conjugate gradient methodology towards the LM algorithm which was studied by Moller [152] and was determined as Scaled Conjugate Algorithm (SCG). The effectiveness of this algorithm depends on the parameters involved in designing the model which is modified at each successive iteration and thus provides a significant benefit compared to algorithms based on line-search.

Researchers from different engineering domains have started adopting the ANN for optimization and classification. The prediction of flash floods being one of the major disasters for the humans depends on different parameters (wave pattern, wind speed, precipitation intensity) and its forecasting on floods with a perfect sensing rate was studied using LM, BR and SCG algorithms [153]. The performance on non-linear data information provided a better result with BR algorithm compared to LM and SCG. In a similar manner, the stock prices in the Indian market which holds a financial data was understood using ANN with three different algorithms [154]. The prices which are dynamic in nature is more difficult for its prediction showed a 99.9% accuracy on the different algorithms from the initial dataset whereas a significant drop of 3 to 4% was observed for data information over a period of 15 min. Amalanathan et al. [155] have studied the impact of ageing on transformer insulation and its classification using Principal component analysis (PCA) and Artificial neural network. From the observations, the authors have concluded that ANN provides a better accuracy towards the interpretation of trained network compared to PCA. Similarly, the ANN regression model have been used for predicting of viscosity of nanofluids using three different algorithms (LM, BR, SCG) and inferred that LM algorithm provided a better prediction than the other techniques [156]. The effect of trash content and truck-related air emissions have been investigated using geographic information system (GIS) along with ANN model [157], and a higher performance was observed only when the maximum values for combined wastes and trash were less prominent in the input data statistics on wastages. Banerjee et al. [158] have understood the wastewater treatment using graphene oxide nanoplatelets with its characterization on toxicity and evaluation on optimal amount of safranin performed using ANN model. From the findings, 99.8% removal of dye was achieved after 2 h of treatment with the safranin solution of 50 mg/L maintained at a pH and temperature of 6.4 and 300 K respectively. Ghosal et al. [159] used the ANN model for optimizing the depth of CO_2 LASER-MIG welding used for alloy containing aluminum and magnesium. The algorithm used backpropagation for training the network using both BR and LM techniques with its optimization results correlating well with the experimental values. Ranade et al. [160] have used the approach of hybrid chemistry to estimate the oxidation of hydrocarbons at high temperatures resulting from experiments on pyrolysis using ANN. The usage of ANN technique for identifying the reaction rates at an early stage led to a reduction in chemical reactions occurring at the pyrolysis stage. The chlorophyll model which is one of the precautionary methods to limit the onset the algal bloom was investigated for its optimization using ANN to reduce the expenses incurred on the marine ecosystem [161]. The biodiesel production from algae oil is gaining more importance in the recent years due to higher amount of oil content and better productivity. Thus, the transesterification of algal oil at low temperatures towards the production of biodiesel was experimented and analyzed using both RSM and ANN [162]. It was inferred that regression analysis provided a better prediction from the ANN model compared to RSM.

2.4 RSM-GA

The response surface methodology (RSM) combines the design of experiments (DOE) along with statistical methods for creating and optimizing empirical models. In the recent years, the search space based on genetic algorithm [163] from the polynomial equation generated from RSM is gaining more importance in the deep learning methodologies. The idea of finding the optimal point in a problem is done after the completion of response surface model where the GA search algorithms is incredibly effective [164]. The integration of RSM with GA for the selection of near optimal target value have been demonstrated by Khoo and Chen in the early twentieth century [165]. The researchers developed an outline on the hybrid prototype model to deal with single and multi-response variables with various constraints.

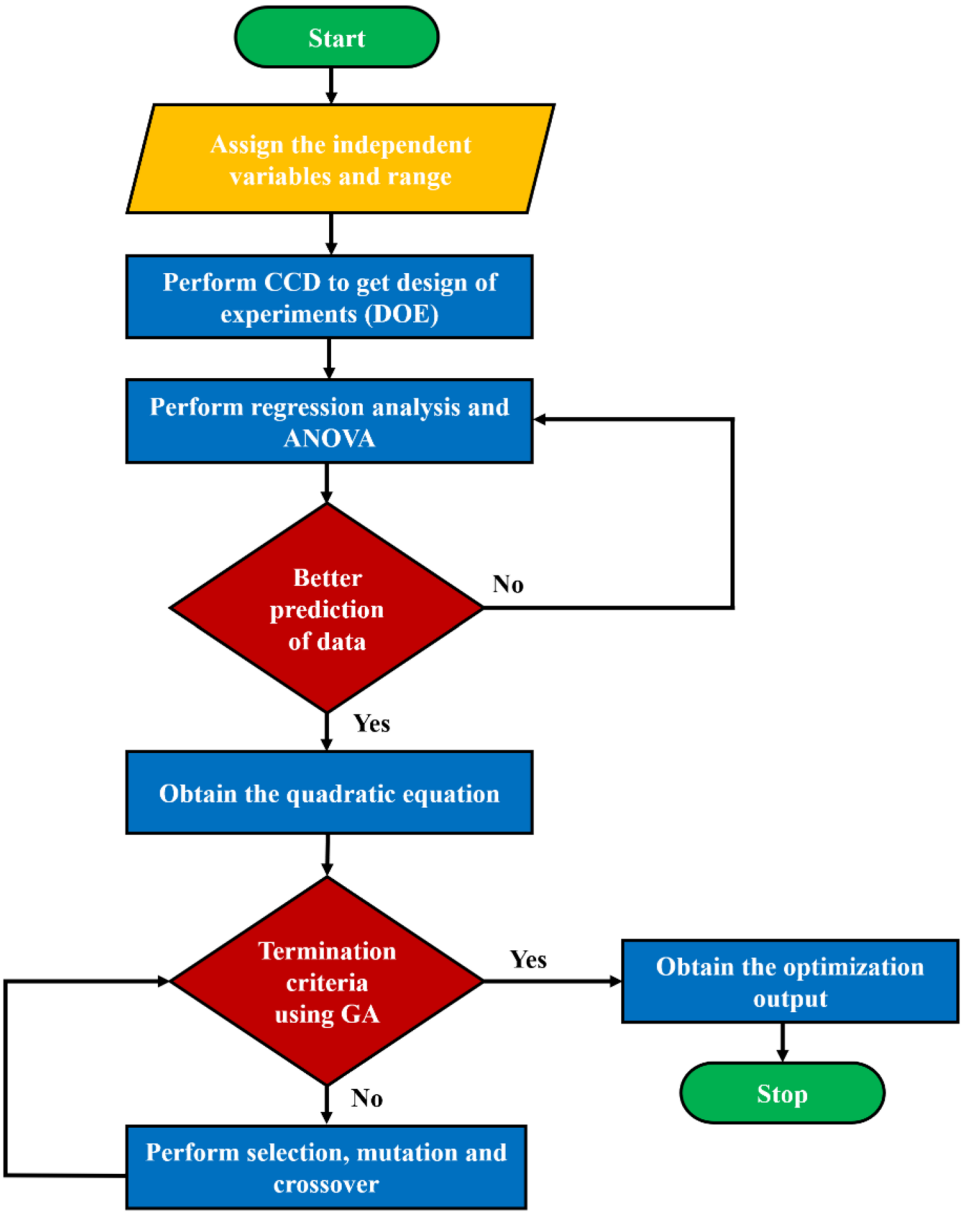


Fig. 7 Flowchart of RSM-GA model

The flowchart indicating the working of RSM-GA model is shown in Fig. 7.

Initially, the independent variables and dependent output response are added to the RSM model with individual range. The design experiments are formulated based on central composite design (CCD) as discussed in Sect. 2.1. Once the experiments are performed, the regression analysis is performed and tested for the analysis of variance (ANOVA) to understand the difference between the predicted and actual results. The quadratic equation depending on the number of variables is generated which is then optimized using GA using suitable parameters such as population size, mutation function, crossover fraction and selection function. The algorithm iterates until it reaches the tolerance limit and maximum number of generations. The isolation of fungus producing proteases from microorganism was assessed with different parameters (temperature, sucrose and pH) to yield a maximum enzyme using hybrid RSM-GA optimization model [166]. The highest enzyme production was observed at pH of 8 with its temperature maintained between 30 and 60 °C. Sabry et al. [167] investigated the impact of tensile strength on aluminum material in frictional stir welding process in underwater considering the diameter, rotating and travelling speed as its input variables. The hybrid model of RSM and GA observed a higher accuracy compared to its individual optimization which improved the issue involved with the welding process in the pipeline. Hasanien et al. [168] developed a design for the cascaded control of power conversion unit using Taguchi method and studied the effect of the design parameters during fault conditions in the grid

using RSM-GA. On comparing both Taguchi method and RSM-GA, the transient responses were found to better in the former compared to later due to larger design experiments. The flocculation process has been induced in microalgae using alkali [169] where the multi-objective optimization gave lower values of input variables yielding a higher efficiency while using RSM model. To the contradict, the RSM-GA model results in higher values for both the input variables and on the efficiency of microalgae. Similarly, the waste water treatment using iron electrode pairs towards the removal of turbidity was experimented [170] with multiple variables such as intensity of current, time for settling and electrolysis process, and temperature. An increase of 3% in the turbidity removal was inferred from hybrid model (RSM-GA) compared to the optimization performed using RSM model separately. The problems involved in the real time applications often involve multiple responses where GA can be well adapted with RSM model in finding the near feasible solutions. Thus, this hybrid combination of RSM and GA provides a better optimization strategy in solving issues for problems involving large number of input variables and output responses.

2.5 ANN-GA

The neural network model developed based on the relation between input and output variables depends on the training, number of hidden layers with multiple trial and error methodologies adopted until a higher prediction in the experimental dataset is being obtained. The coupling of ANN with GA was introduced in order to attain a higher precision in finding the optimum point from the search space [171, 172]. The ANN trained using a feedforward or backpropagation algorithm is used to identify the fitness function which coupled with GA formulates the objective function towards optimization problem. It is possible to create the best ANN model for usage in a specific problem using a variety of GA techniques. GA is used to improve interpretation, topology, feature selection, training, and weights associated with ANN [173]. The major problem involved with ANN is deciding the optimal configurations required for the network such as number of layers, neurons in the hidden layer and activation functions. There is no clear methodology used for architecture settings involved in ANN where its coupling with GA can create a better optimal design thereby improving its reliability and performance of the network [174]. It is well known that improvement in the ANN model is achieved in finding its optimal weights. The ANN coupled with GA can be used to identify the optimum weights where the probability of termination towards local minima using gradient descent method is overcome with its convergence to global minimum value [175]. Further, selecting a suitable input dataset is an important issue with ANN model where its coupling with GA identifies the required input dataset reducing their dimensional space and statistically improving their selection accuracy better than traditional methods [176]. Thus, ANN performs well in coherence with GA in finding the optimal model and approximating the parameters to increase their efficiency. Figure 8 shows the steps involved in formulating the ANN-GA model. Initially, the ANN model creates the input layer and output layer with optimized number of hidden layers chosen and trained with a suitable algorithm as mentioned in Sect. 2.3.

The model is allowed for its training until a higher R^2 value is obtained. If a very lower precision is obtained, then network is trained again with a modification of weights and bias in a suitable neuron of hidden layer and output layer [177]. Once a better prediction of network is being created with lower mean square error, the weights, bias and activation functions used for transformations are formulated to obtain the required objective function. This functional equation is then optimized using GA with a suitable constraints (lower bounds, upper bounds, nonlinear constraints) and algorithm settings [178]. The iteration is continued until a lower tolerance in finding the optimal point is being reached else the GA parameters (selection, mutation and crossover) are to be reverted back for its modifications. Bahrami et al. [179] investigated the design towards the inflow of groundwater using hybrid coupling methods ANN-GA and simulated annealing techniques. Among the different hybrid algorithms, the ANN-GA observed a better correlation which was helpful for mining engineers to approach an effective management towards the controlling of water in mining. The starting of combustion in engines considering the mixture of air and fuel along with ignition timing was trained with neural network and its optimization results using GA technique enhanced the performance of neural network [180]. The coal fired power plant contributes a major part in the generation of electrical energy throughout the world. The parameters affecting the power generation (pressure, temperature, air ratio for fuel) was optimized using ANN-GA method [181] and a better plant efficiency was obtained due to the reduction of fuel consumption. The transformer which is responsible for reliable electrical supply from the generation towards distribution unit involves different faults during its operational life time. The identification of faults based on the dissolved gas performance and ANN-GA modelling has provided a better classification of faults which is helpful for insulation engineers [182]. Bülbül et al. [183] determined the risk involved with reinforced concrete buildings from the database on 329 buildings in Bitlis, Turkey with the hybrid coupling of ANN-GA. The initial population of GA was performed as the first step where its initial parameters (population number, selection and mutation rate, iterations) govern the network

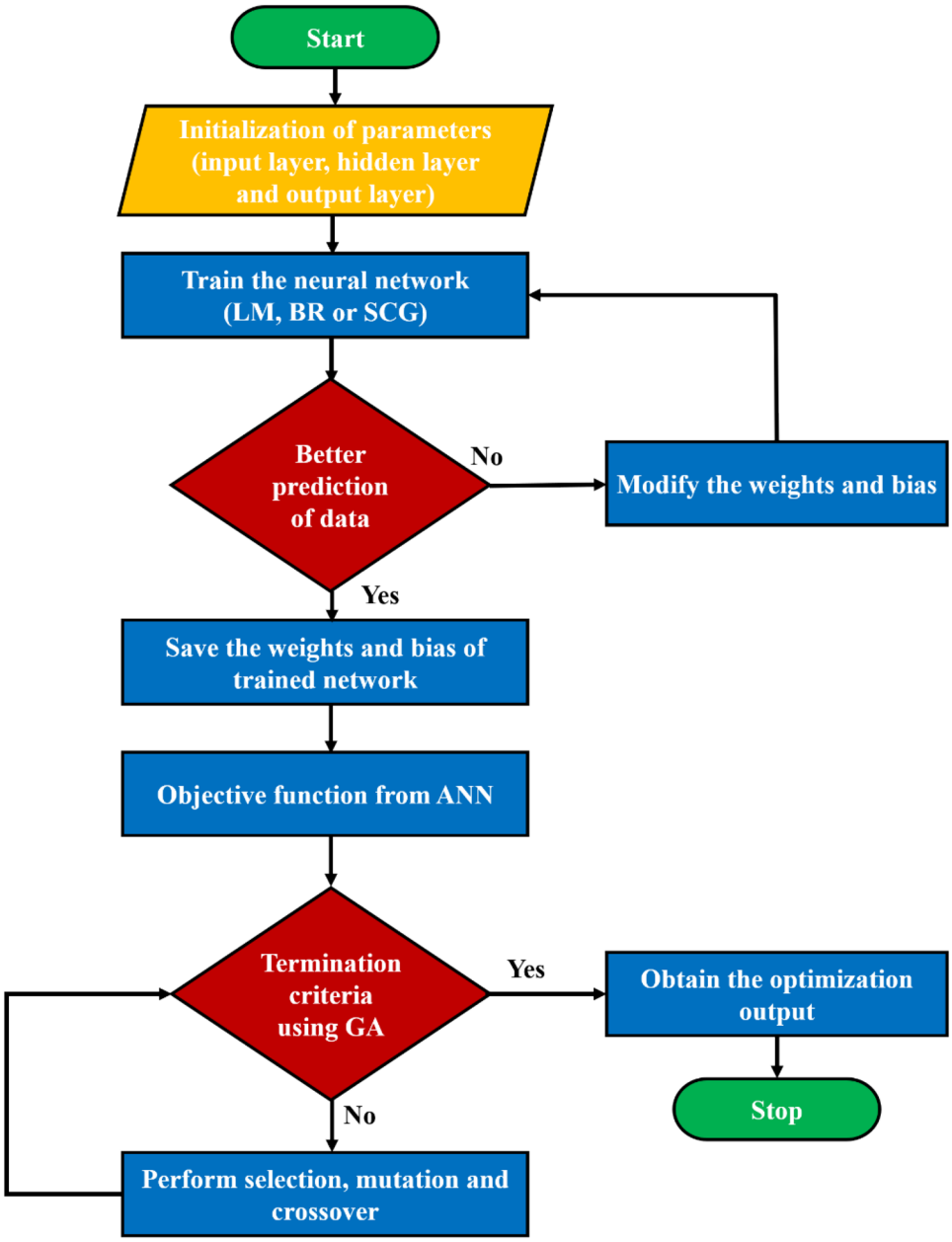


Fig. 8 Flowchart of ANN-GA model

parameters (number of input layers, hidden layers, activation, and training function) of ANN. The fitness value of each gene used for generating the ANN structure was performed for selection, cross-over and mutation process to identify the most successful gene towards the hybrid model. The proposed hybrid model provided a better network parameter (98% accuracy) in identifying the earthquake risks involved in RC buildings which is not possible with the traditional trial and error methodology. Smaali et al. [184] experimented the degradation of Azithromycin (AZM) which was considered as one of the major drug used during the pandemic situation of COVID 19. The Like-Fenton experiment was used for identifying the AZM degradation rate considering the impact of various factors such as pH, initial concentrations, doses of $FeSO_4$ and NaClO). The optimization process was performed using the ANN-GA algorithm to identify the conditions leading to maximum AZM degradation. The ANN-based model was first determined with additional responses developed from central composite design (CCD) for higher accuracy which then coupled with GA determined the conditions towards maximum AZM degradation rate. Thus, the nonlinear regression analysis provided a better model with ANN-GA algorithm which could be a major alternative towards pharmaceutical industry. The nanotechnology that has got its applications towards various engineering

fields, where the prediction on the density of nanofluids with respect to several parameters (temperature, volume fraction, density of base fluid) is now possible using hybrid ANN-GA model [185]. Thus, the hybrid coupling of ANN-GA model could provide a better optimization result compared to individual algorithm models. Nevertheless, the training of ANN towards the dataset can impact the optimization problem performed with GA. Hence, a suitable model framed through the ANN is a precautionary measure to be followed in this hybridization which should be considered by the deep learning researchers before performing the optimization.

3 Future Perspective and Conclusions

Optimization is one of the key parameters required in the different applications of engineering domains. The techniques used for implementing the optimization (RSM, GA and ANN) can result in different accuracies based on the problems involved within the search space. Despite being effective, capable of handling issues involving numerous design variables, taking interaction effects into account, and requiring minimum parameter modification, RSM only provides local optimal solutions. The problem of creating the objecting function along with multiple operators in GA makes them computationally complex. Further, the selection method used for fitness function evaluation should yield a result without premature convergence. The optimal design of ANN model with required neurons in the hidden layer and appropriate activation function is tedious process requiring trial and error methods with its training and testing to be validated with multiple algorithms. Thus, using a technique called hybridization, which combines two separate algorithms (RSM-GA and ANN-GA), it is possible to overcome the above limitations involved with the individual optimization methods. In addition, the hybrid statistical approach is known to increase the accuracy of the process variables and responses in detecting the global optima involved in the optimization process compared to traditional single techniques. Hybrid statistical approach is known to increase the accuracy of process variables and responses when compared to traditional single techniques. Thus, these optimization strategies could be useful for the engineers working in the various technical fields and may pave the way for real time applications.

Author Contributions AS-Conceptualization, methodology, writing original draft; PM-Review and Supervision; AAJ-Conceptualization, methodology, writing original draft. All the authors have read and agreed to the published version of the manuscript.

Funding The authors did not receive any specific grant from funding agencies in public or non-profit sector.

Data Availability The datasets generated during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare no conflict of interest.

Ethical Approval The authors of the manuscript have fulfilled with all the ethical standards necessary for the publication of the journal.

Replication of Results No results are presented.

References

1. Onwubolu GC, Babu BV (2004) New optimization techniques in engineering. Springer Berlin Heidelberg, Berlin
2. Pistikopoulos EN et al (2021) Process systems engineering—the generation next? Comput Chem Eng 147:107252. https://doi.org/10.1016/j.compchemeng.2021.107252
3. Strawderman WE (2002) Practical optimization methods with mathematica (R) applications, optimization foundations and applications. J Am Stat Assoc 97(457):366–366. https://doi.org/10.1198/jasa.2002.s467
4. Sun S, Cao Z, Zhu H, Zhao J (2020) A survey of optimization methods from a machine learning perspective. IEEE Trans Cybern 50(8):3668–3681. https://doi.org/10.1109/TCYB.2019.2950779
5. Spedicato E, Xia Z, Zhang L (2008) ABS algorithms for optimization. Encycl Optim. https://doi.org/10.1007/978-0-387-74759-0_2
6. Arora RK (2016) Optimization: algorithms and applications. Choice Rev Online. https://doi.org/10.5860/choice.195857
7. Wilson DR, Martinez TR (2003) The general inefficiency of batch training for gradient descent learning. Neural Netw 16(10):1429–1451. https://doi.org/10.1016/S0893-6080(03)00138-2
8. Sergeyev YD, Kvasov DE, Mukhametzhanov MS (2018) On the efficiency of nature-inspired metaheuristics in expensive global optimization with limited budget. Sci Rep 8(1):1–9. https://doi.org/10.1038/s41598-017-18940-4
9. Yaqoob I et al (2016) Big data: from beginning to future. Int J Inf Manag 36(6):1231–1247. https://doi.org/10.1016/j.ijinfomgt.2016.07.009
10. Dean A, Voss D, Draguljić D (2017) Response surface methodology. Springer, Cham, pp 565–614
11. Wong WK, Ming CI (2019) A review on metaheuristic algorithms: recent trends, benchmarking and applications. Int Conf Smart Comput Commun ICSCC 2019:1–5. https://doi.org/10.1109/ICSCC.2019.8843624
12. Dillen W, Lombaert G, Schevenels M (2021) Performance assessment of metaheuristic algorithms for structural optimization taking into account the influence of algorithmic control parameters. Front Built Environ 7(March):1–16. https://doi.org/10.3389/fbuil.2021.618851
13. Agrawal P, Abutarboush HF, Ganesh T, Mohamed AW (2021) Metaheuristic algorithms on feature selection: a survey of one decade of research (2009–2019). IEEE Access 9:26766–26791. https://doi.org/10.1109/ACCESS.2021.3056407

14. Said GAEA, Mahmoud AM (2014) A comparative study of meta-heuristic algorithms for solving quadratic assignment problem. Int J Adv Comput Sci Appl 5(1):1–6
15. Abdel-Basset M, Abdel-Fatah L, Sangaiah AK (2018) Chapter 10—Metaheuristic algorithms: a comprehensive review. Computational intelligence for multimedia big data on the cloud with engineering applications. Elsevier, Amsterdam, pp 185–231
16. Zhang G, Pan L, Neri F, Gong M, Leporati A (2017) Metaheuristic optimization: algorithmic design and applications. J Optim 2017:2–4
17. Hussain K, Najib M, Salleh M, Cheng S, Shi Y (2019) Metaheuristic research: a comprehensive survey. Artif Intell Rev 52(4):2191–2233. https://doi.org/10.1007/s10462-017-9605-z
18. Slowik A, Kwasnicka H (2020) Evolutionary algorithms and their applications to engineering problems. Neural Comput Appl 32(16):12363–12379. https://doi.org/10.1007/s00521-020-04832-8
19. Jangid S, Puri R (2019) Evolutionary algorithms: a critical review and its future prospects. Int J Trend Res Dev 1:7–9
20. Chakraborty A, Kar AK (2017) Swarm intelligence: a review of algorithms swarm intelligence: a review of algorithms. Model Optim Sci Technol. https://doi.org/10.1007/978-3-319-50920-4
21. Brezocnik L, Fister JI, Podgorelec V (2018) Swarm intelligence algorithms for feature selection. Appl Sci. https://doi.org/10.3390/app8091521
22. Singh A, Kumar A (2021) Applications of nature-inspired meta-heuristic algorithms: a survey applications of nature-inspired meta-heuristic algorithms: a survey Avjeet Singh* and Anoj Kumar. Int J Adv Intell Paradig. https://doi.org/10.1504/IJAIP.2021.10027703
23. Priyadarshini J, Premalatha M, Cep R, Jayasudha M, Kalita K (2023) Analyzing physics-inspired metaheuristic algorithms in feature selection with K-nearest-neighbor. Appl Sci 13:906
24. Biswas A, Mishra KK, Tiwari S, Misra AK (2013) Physics-inspired optimization algorithms: a survey. J Optim 2013:1
25. Beheshti Z, Mariyam S, Shamsuddin H (2013) A review of population-based meta-heuristic algorithm. Int J Adv Soft Comput Appl 5(1):1–35
26. Rai R, Das A, Ray S, Gopal K (2022) Human—inspired optimization algorithms: theoretical foundations, algorithms, open—research issues and application for multi—level thresholding. Springer, Netherlands
27. Mingyi Zhang Y, Zhang Y (2013) The human-inspired algorithm: a hybrid nature-inspired approach to optimizing continuous functions with constraints. J Comput Intell Electron Syst 2(1):80–87
28. Kumar V (2022) A state-of-the-art review of heuristic and metaheuristic optimization techniques for the management of water resources. Water supply 22(4):3702–3728. https://doi.org/10.2166/ws.2022.010
29. Bhosale V, Shastri SS, Khandare A (2017) A review of genetic algorithm used for optimizing scheduling of resource constraint construction projects. Int Res J Eng Technol 1:2869–2872
30. Sioshansi R, Conejo AJ (2017) Optimization in engineering, 1st edn. Springer International Publishing, Cham
31. Rao SS (2019) Engineering optimization: theory and practice. John Wiley & Sons
32. Chen Wu Y, Wen Feng J (2018) Development and application of artificial neural network. Wirel Pers Commun 102(2):1645–1656. https://doi.org/10.1007/s11277-017-5224-x
33. Ilin R, Kozma R, Werbos PJ (2008) Beyond feedforward models trained by backpropagation: a practical training tool for a more efficient universal approximator. IEEE Trans Neural Networks 19(6):929–937. https://doi.org/10.1109/TNN.2008.2000396
34. Apicella A, Donnarumma F, Isgrò F, Prevete R (2021) A survey on modern trainable activation functions. Neural Netw 138:14–32. https://doi.org/10.1016/j.neunet.2021.01.026
35. Bhattacharya SS, Garlapati VK, Banerjee R (2011) Optimization of laccase production using response surface methodology coupled with differential evolution. N Biotechnol 28(1):31–39. https://doi.org/10.1016/j.nbt.2010.06.001
36. Sarker IH (2021) Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. SN Comput Sci 2(6):1–20. https://doi.org/10.1007/s42979-021-00815-1
37. Jha AK, Sit N (2021) Comparison of response surface methodology (RSM) and artificial neural network (ANN) modelling for supercritical fluid extraction of phytochemicals from Terminalia chebula pulp and optimization using RSM coupled with desirability function (DF) and genetic. Ind Crops Prod 170:113769. https://doi.org/10.1016/j.indcrop.2021.113769
38. Deshwal S, Kumar A, Chhabra D (2020) Exercising hybrid statistical tools GA-RSM, GA-ANN and GA-ANFIS to optimize FDM process parameters for tensile strength improvement. CIRP J Manuf Sci Technol 31:189–199. https://doi.org/10.1016/j.cirpj.2020.05.009
39. Kumari M, Gupta SK (2019) Response surface methodological (RSM) approach for optimizing the removal of trihalomethanes (THMs) and its precursor's by surfactant modified magnetic nanoadsorbents (sMNP)—an endeavor to diminish probable cancer risk. Sci Rep 9(1):1–11. https://doi.org/10.1038/s41598-019-54902-8
40. Sarabia LA, Ortiz MC (2009) Response surface methodology. Comprehensive chemometrics. Elsevier, Amsterdam, pp 345–390
41. M. C. Fu, *Handbook of simulation optimization*, vol. 216. 2015.
42. Khuri S, Mukhopadhyay AI (2010) Response surface methodology. Wiley Interdiscip Rev Comput Stat 2(2):128–149
43. Draper DK, Lin NR (1996) Response surface designs. Design Analy Exp 13:343–375
44. Dean D, Voss A, Draguljić D, Dean D, Voss A, Draguljić D (2017) Response surface methodology. Des Anal Exp. https://doi.org/10.1007/978-3-319-52250-0_16
45. Sarabia LA, Ortiz MC (2009) Response surface methodology. Compr Chemom 1(2):345–390. https://doi.org/10.1016/B978-044452701-1.00083-1
46. Myers S, Vining RH, Giovannitti-Jensen GG, Myers A (1992) Variance dispersion properties of second-order response surface designs. J Qual Technol 24(1):1–11
47. Hadiyat MA, Sopha BM, Wibowo BS (2022) Response surface methodology using observational data: a systematic literature review. Appl Sci. https://doi.org/10.3390/app122010663
48. Hanrahan D, Zhu G, Gibani J, Patil DG (2005) Chemometrics and statisticslexperimental design. Encycl Anal Sci 8:13
49. Ismail M, Author C (2013) Alternative approach to fitting first-order model to the response surface methodology. Pakistan J Commer Soc Sci 7(1):157–165
50. Lamidi I, Olaleye S, Bankole N, Obalola Y, Aribike A, Adigun I (2022) Applications of response surface methodology (RSM) in product design, development, and process optimization. Response Surf Methodol Res Adv Appl. https://doi.org/10.5772/intechopen.106763
51. Alrweili H, Georgiou S, Stylianou S (2020) A new class of second-order response surface designs. IEEE Access 8:115123–115132. https://doi.org/10.1109/ACCESS.2020.3001621
52. Gunawan A (2014) Institutional knowledge at Singapore management university second order-response surface model for the automated parameter tuning problem second order-response

surface model for the automated parameter tuning problem. IEEE Int Conf Ind Eng Eng Manag 2014:652–656
53. Saini N (2017) Review of selection methods in genetic algorithms. Int J Eng Comput Sci 6(12):23261–23263. https://doi.org/10.18535/ijecs/v6i12.04
54. Dasgupta T, Pillai NS, Rubin DB (2015) Causal inference from 2 K factorial designs by using potential outcomes. J R Stat Soc Ser B 77(4):727–753
55. Branson Z, Dasgupta T, Rubin DB (2016) Improving covariate balance in 2k factorial designs via rerandomization with an application to a New York city department of education high school study. Ann Appl Stat 10(4):1958–1976. https://doi.org/10.1214/16-AOAS959
56. Kandananond K (2013) Applying 2k factorial design to assess the performance of ANN and SVM methods for forecasting stationary and non-stationary time series. Procedia Comput Sci 22:60–69. https://doi.org/10.1016/j.procs.2013.09.081
57. Tablets EMI (2021) Application of Plackett–Burman design of experiments in the identification of main factors' in the formulation of dabigatran etexilate mesylate immediate-release tablets. Int J Pharm Sci Res 12(12):6587–6592. https://doi.org/10.13040/IJPSR.0975-8232.12(12).6587-92
58. Ekpenyong MG, Antai SP, Asitok D, Ekpo BO (2017) Plackett–Burman design and response surface optimization of medium trace nutrients for glycolipopeptide biosurfactant production. Iran Biomed J 21(4):249–260. https://doi.org/10.18869/acadpub.ibj.21.4.249
59. Chaudhari SR (2020) Application of Plackett–Burman and central composite designs for screening and optimization of factor influencing the chromatographic conditions of HPTLC method for quantification of efonidipine hydrochloride. J Anal Sci Technol 11(48):1–13
60. Peele A, Krupanidhi S, Reddy ER, Indira M, Bobby N (2018) Plackett–Burman design for screening of process components and their effects on production of lactase by newly isolated *Bacillus* sp. VUVD101 strain from Dairy effluent. Beni-Suef Univ J Basic Appl Sci 7(4):543–546. https://doi.org/10.1016/j.bjbas.2018.06.006
61. Patel MB, Shaikh F, Patel V, Surti NI (2017) Application of simplex centroid design in formulation and optimization of floating matrix tablets of metformin. J Appl Pharm Sci 7(4):23–30. https://doi.org/10.7324/JAPS.2017.70403
62. Bahramparvar M, Tehrani MM (2015) Application of simplex-centroid mixture design to optimize stabilizer combinations for ice cream manufacture. J Food Sci Technol 52(3):1480–1488. https://doi.org/10.1007/s13197-013-1133-5
63. Article R, Reji M, Kumar R (2023) Response surface methodology (RSM): an overview to analyze multivariate data. Indian J Microbiol Res 9(4):241–248
64. Phanphet S (2021) Application of full factorial design for optimization of production process by turning machine. J Tianjin Univ Sci Technol ISSN 54(08):35–55. https://doi.org/10.17605/OSF.IO/3TESD
65. Al Sadi J (2018) Designing experiments: 3 level full factorial design and variation of processing parameters methods for polymer colors. Adv Sci Technol Eng Syst J 3(5):109–115
66. Salihu MM, Nwaosu CS (2021) Discrimination between 2k and 3k factorial designs using optimality based criterion Murtala Muhammad Salihu and Chigozie Sylvester Nwaosu. African Sch J pure Appl Sci 22(9):79–94
67. Aggarwal ML, Kaul R (1999) Hidden projection properties of some optimal designs. Stat Probab Lett 43(1):87–92. https://doi.org/10.1016/S0167-7152(98)00249-1
68. Kasina MM, Joseph K, John M (2020) Application of central composite design to optimize spawns propagation. Open J Optim 9:47–70. https://doi.org/10.4236/ojop.2020.93005
69. Sadhukhan B, Mondal NK, Chattoraj S (2016) ScienceDirect optimisation using central composite design (CCD) and the desirability function for sorption of methylene blue from aqueous solution onto Lemna major. Karbala Int J Mod Sci 2(3):145–155. https://doi.org/10.1016/j.kijoms.2016.03.005
70. Bayuo J, Abdullai M, Kenneth A, Pelig B (2020) Optimization using central composite design (CCD) of response surface methodology (RSM) for biosorption of hexavalent chromium from aqueous media. Appl Water Sci 10(6):1–12. https://doi.org/10.1007/s13201-020-01213-3
71. Hassan H, Adam SK, Alias E, Mohd M, Meor R, Affandi M (2021) Central composite design for formulation and optimization of solid lipid nanoparticles to enhance oral bioavailability of acyclovir. Molecules 26(5432):1–19
72. Alam P et al (2022) Box–Behnken design (BBD) application for optimization of chromatographic conditions in RP-HPLC method development for the estimation of thymoquinone in *Nigella sativa* seed powder. Processes 10:1082
73. Yadav P, Rastogi V, Verma A (2020) Application of Box–Behnken design and desirability function in the development and optimization of self-nanoemulsifying drug delivery system for enhanced dissolution of ezetimibe. Futur J Pharm Sci 6(1):1–20
74. Darvishmotevalli M, Zarei A, Moradnia M, Noorisepehr M, Mohammadi H (2019) Optimization of saline wastewater treatment using electrochemical oxidation process: Prediction by RSM method. MethodsX 6:1101–1113. https://doi.org/10.1016/j.mex.2019.03.015
75. Singh Pali H, Sharma A, Kumar N, Singh Y (2021) Biodiesel yield and properties optimization from Kusum oil by RSM. Fuel 291:120218. https://doi.org/10.1016/j.fuel.2021.120218
76. Behera SK, Meena H, Chakraborty S, Meikap BC (2018) Application of response surface methodology (RSM) for optimization of leaching parameters for ash reduction from low-grade coal. Int J Min Sci Technol 28(4):621–629. https://doi.org/10.1016/j.ijmst.2018.04.014
77. Hatami M (2017) Nanoparticles migration around the heated cylinder during the RSM optimization of a wavy-wall enclosure. Adv Powder Technol 28(3):890–899. https://doi.org/10.1016/j.apt.2016.12.015
78. Raj RA, Murugesan S (2022) Optimization of dielectric properties of pongamia pinnata methyl ester for power transformers using response surface methodology. IEEE Trans Dielectr Electr Insul 29(5):1931–1939. https://doi.org/10.1109/TDEI.2022.3190257
79. Kinnear KE (1994) A perspective on the work in this book. In: Kinnear KE (ed) Advances in genetic programming. MIT Press, pp 3–17
80. Carr J (2014) An introduction to genetic algorithms. Senior Project 1(40):7
81. Forrest S (1996) Genetic algorithms. ACM Comput Surv. https://doi.org/10.1145/234313.234350
82. McCall J (2005) Genetic algorithms for modelling and optimisation. J Comput Appl Math 184(1):205–222. https://doi.org/10.1016/j.cam.2004.07.034
83. Lambert-torres G, Martins HG, Coutinho MP, Silva LEB, Matsunaga FM, Carminati RA (2009) Genetic algorithm to system restoration. World Congress Electron Electric Eng. https://doi.org/10.13140/RG.2.1.4926.2482
84. Lima AR, de Mattos Neto PSG, Silva DA, Ferreira TAE (2016) Tests with different fitness functions for tuning of artificial neural networks with genetic algorithms. X Congresso Brasileiro de Inteligˆencia Computacional. 1(1):1–8. https://doi.org/10.21528/cbic2011-32.5
85. Kour H, Sharma P, Abrol P (2015) Analysis of fitness function in genetic algorithms. J Sci Tech Adv 1(3):87–89

86. Mandal S, Anderson TA, Turek JS, Gottschlich J, Zhou S, Muzahid A (2021) Learning fitness functions for machine programming. Proc Mach Learn Syst 1:139–155
87. Petridis V, Kazarlis S, Bakirtzis A (1998) Varying fitness functions in genetic algorithm constrained optimization: The cutting stock and unit commitment problems. IEEE Trans Syst Man Cybern Part B Cybern 28(5):629–640. https://doi.org/10.1109/3477.718514
88. Avdeenko TV, Serdyukov KE, Tsydenov ZB (2021) Formulation and research of new fitness function in the genetic algorithm for maximum code coverage. Procedia Comput Sci 186:713–720. https://doi.org/10.1016/j.procs.2021.04.194
89. Der Yang M, Yang YF, Su TC, Huang KS (2014) An efficient fitness function in genetic algorithm classifier for landuse recognition on satellite images. Sci World J. https://doi.org/10.1155/2014/264512
90. Chakraborty B (2002) Genetic algorithm with fuzzy fitness function for feature selection. IEEE Int Symp Ind Electron 1:315–319. https://doi.org/10.1109/isie.2002.1026085
91. Büche D, Schraudolph NN, Koumoutsakos P (2005) Accelerating evolutionary algorithms with Gaussian process fitness function models. IEEE Trans Syst Man Cybern Part C Appl Rev 35(2):183–194. https://doi.org/10.1109/TSMCC.2004.841917
92. Katoch S, Chauhan SS, Kumar V (2021) A review on genetic algorithm: past, present, and future. Multimed Tools Appl. https://doi.org/10.1007/s11042-020-10139-6
93. Haq EU, Ahmad I, Hussain A, Almanjahie IM (2019) A novel selection approach for genetic algorithms for global optimization of multimodal continuous functions. Comput Intell Neurosci. https://doi.org/10.1155/2019/8640218
94. Shukla A, Pandey HM, Mehrotra D (2015) Comparative review of selection techniques in genetic algorithm. Int Conf Futur Trends Comput Anal Knowl Manag. https://doi.org/10.1109/ABLAZE.2015.7154916
95. Jebari K, Madiafi M (2013) Selection methods for genetic algorithms. Int J Emerg Sci 3(4):333–344
96. Minetti G, Salto C, Alfonso H (1999) A study of performance of stochastic universal sampling versus proportional selection on genetic algorithms, I work. Investig en Ciencias la Comput 1:9–12
97. Pencheva T, Atanassov K, Shannon A (2009) Modelling of a stochastic universal sampling selection operator in genetic algorithms using generalized nets. Tenth Int Work Gen Nets 2009:1–7
98. Orong MY, Sison AM, Hernandez AA (2018) Mitigating vulnerabilities through forecasting and crime trend analysis. Eng Bus Soc Sci. https://doi.org/10.1109/ICBIR.2018.8391166
99. Hancock PJB (2019) Selection methods for evolutionary algorithms. Practical handbook of genetic algorithms. CRC Press, Boca Raton, pp 67–92
100. Champlin R, Champlin R (2018) Selection methods of genetic algorithms selection methods of genetic algorithms. Digit Commons Comput Sci 8:1
101. Kiran CA, Xaxa D (2015) Comparative study on various selection methods in genetic algorithm. Int J Soft Comput Artif Intell 8(3):96–103
102. Jannoud I, Jaradat Y, Masoud MZ, Manasrah A, Alia M (2022) The role of genetic algorithm selection operators in extending wsn stability period: a comparative study. Electron 11(1):1–16. https://doi.org/10.3390/electronics11010028
103. Umbarkar AJ, Sheth PD (2015) Crossover operators in genetic algorithms: a review. ICTACT J Soft Comput 06(01):1083–1092. https://doi.org/10.21917/ijsc.2015.0150
104. Y. Kaya, M. Uyar, and R. Tek, "A novel crossover operator for genetic algorithms: ring crossover," *arXiv Prepr.*, pp. 1–4, 2011.
105. Anand E, Panneerselvam R (2016) A study of crossover operators for genetic algorithm and proposal of a new crossover operator to solve open shop scheduling problem. Am J Ind Bus Manag 06(06):774–789. https://doi.org/10.4236/ajibm.2016.66071
106. Kora P, Yadlapalli P (2017) Crossover operators in genetic algorithms: a review. Int J Comput Appl 162(10):34–36. https://doi.org/10.5120/ijca2017913370
107. Magalhães-Mendes J (2013) A comparative study of crossover operators for genetic algorithms to solve the job shop scheduling problem. WSEAS Trans Comput 12(4):164–173
108. Kumar VS, Panneerselvam R (2017) A study of crossover operators for genetic algorithms to solve VRP and its variants and new sinusoidal motion crossover operator. Int J Comput Intell Res 13(7):1717–1733
109. Hassanat A, Almohammadi K, Alkafaween E, Abunawas E, Hammouri A, Prasath VBS (2019) Choosing mutation and crossover ratios for genetic algorithms-a review with a new dynamic approach. Information. https://doi.org/10.3390/info10120390
110. O. Abdoun, J. Abouchabaka, and C. Tajani, "Analyzing the performance of mutation operators to solve the travelling salesman problem," *arXiv Prepr*, 2012.
111. Lim SM, Sultan ABM, Sulaiman MN, Mustapha A, Leong KY (2017) Crossover and mutation operators of genetic algorithms. Int J Mach Learn Comput 7(1):9–12. https://doi.org/10.18178/ijmlc.2017.7.1.611
112. Nazeri Z, Khanli LM (2014) An insertion mutation operator for solving project scheduling problem. Iran Conf Intell Syst ICIS 2014:1–4. https://doi.org/10.1109/IranianCIS.2014.6802537
113. Soni N, Kumar T (2014) Study of various mutation operators in genetic algorithms. Int J Comput Sci Inf Technol 5(3):4519–4521
114. Liu C, Kroll A (2016) Performance impact of mutation operators of a subpopulation-based genetic algorithm for multi-robot task allocation problems. Springerplus. https://doi.org/10.1186/s40064-016-3027-2
115. Deep K, Mebrahtu H (2011) Combined mutation operators of genetic algorithm for the travelling salesman problem Kusum. Int J Comb Optim Probl Inform 2(3):1–23
116. Sutton AM, Whitley LD (2014) Fitness probability distribution of bit-flip mutation. Evol Comput 23(2):217–248. https://doi.org/10.1162/EVCO
117. De Falco I, Della Cioppa A, Tarantino E (2002) Mutation-based genetic algorithm: performance evaluation. Appl Soft Comput 1(4):285–299. https://doi.org/10.1016/S1568-4946(02)00021-2
118. S. Sarmady, "An Investigation on Genetic Algorithm Parameters," 2007.
119. Kumar R, Memoria M, Chandel A (2020) Performance analysis of proposed mutation operator of genetic algorithm under scheduling problem. Proc Int Conf Intell Eng Manag ICIEM 2020:193–197. https://doi.org/10.1109/ICIEM48762.2020.9160215
120. W et al (2020) Metadata of the chapter that will be visualized in OnlineFirst. Itib. https://doi.org/10.1007/978-3-642-03503-6
121. Contras D, Matei O (2016) Translation of the mutation operator from genetic algorithms to evolutionary ontologies. Int J Adv Comput Sci Appl 7(1):3–8. https://doi.org/10.14569/ijacsa.2016.070186
122. Bajpai P, Kumar M (2010) Genetic algorithm–an approach to solve global optimization problems. Indian J Comput Sci Eng 1(3):199–206
123. Elsayed SM, Sarker RA, Essam DL (2014) A new genetic algorithm for solving optimization problems. Eng Appl Artif Intell 27:57–69. https://doi.org/10.1016/j.engappai.2013.09.013

124. Jin Y-F, Yin Z-Y, Shen S-L, Hicher P-Y (2016) Selection of sand models and identification of parameters using an enhanced genetic algorithm. Int J Numer Anal Methods Geomech 40(8):1219–1240. https://doi.org/10.1002/nag.2487
125. Marcovecchio MG, Aguirre PA, Scenna NJ (2005) Global optimal design of reverse osmosis networks for seawater desalination: modeling and algorithm. Desalination 184(1–3):259–271. https://doi.org/10.1016/j.desal.2005.03.056
126. Al-Obaidi MA, Li JP, Kara-Zaïtri C, Mujtaba IM (2017) Optimisation of reverse osmosis based wastewater treatment system for the removal of chlorophenol using genetic algorithms. Chem Eng J 316:91–100. https://doi.org/10.1016/j.cej.2016.12.096
127. Ho SH, Chen CY, Lee DJ, Chang JS (2011) Perspectives on microalgal CO_2-emission mitigation systems—a review. Biotechnol Adv 29(2):189–198. https://doi.org/10.1016/j.biotechadv.2010.11.001
128. Azari A, Tavakoli H, Barkdoll BD, Haddad OB (2020) Predictive model of algal biofuel production based on experimental data. Algal Res 47:101843. https://doi.org/10.1016/j.algal.2020.101843
129. Al-Turjman F, Abujubbeh M (2019) IoT-enabled smart grid via SM: an overview. Futur Gener Comput Syst 96:579–590. https://doi.org/10.1016/j.future.2019.02.012
130. Arabali A, Ghofrani M, Etezadi-Amoli M, Fadali MS, Baghzouz Y (2013) Genetic-algorithm-based optimization approach for energy management. IEEE Trans Power Deliv 28(1):162–170. https://doi.org/10.1109/TPWRD.2012.2219598
131. Rosa JPS, Guerra DJD, Horta NCG, Martins RMF, Lourenço NCC (2020) Overview of artificial neural networks. Appl Sci Technol 1:21–44. https://doi.org/10.1007/978-3-030-35743-6_3
132. Babu MS, Imai T, Sarathi R (2021) Classification of aged epoxy micro-nanocomposites through PCA- and ANN-Adopted LIBS analysis. IEEE Trans Plasma Sci 49(3):1088–1096. https://doi.org/10.1109/TPS.2021.3061410
133. Raghavendra U, Fujita H, Bhandary SV, Gudigar A, Tan JH, Acharya UR (2018) Deep convolution neural network for accurate diagnosis of glaucoma using digital fundus images. Inf Sci 441:41–49. https://doi.org/10.1016/j.ins.2018.01.051
134. Ghimire S, Deo RC, Downs NJ, Raj N (2019) Global solar radiation prediction by ANN integrated with European centre for medium range weather forecast fields in solar rich cities of Queensland Australia. J Clean Prod 216:288–310. https://doi.org/10.1016/j.jclepro.2019.01.158
135. Ertuğrul ÖF (2018) A novel type of activation function in artificial neural networks: trained activation function. Neural Netw 99:148–157. https://doi.org/10.1016/j.neunet.2018.01.007
136. Montesinos López OA, Montesinos López A, Crossa J (2022) Multivariate statistical machine learning methods for genomic prediction. Springer, Cham
137. Jain AK, Mao J, Mohiuddin KM (1996) Artificial neural networks: a tutorial. Computer (Long Beach Calif) 29(3):31–44. https://doi.org/10.1109/2.485891
138. Feng J, Lu S (2019) Performance analysis of various activation functions in artificial neural networks. J Phys Conf Ser 1237(2):111–122. https://doi.org/10.1088/1742-6596/1237/2/022030
139. Sharma S, Sharma S, Athaiya A (2020) Activation functions in neural networks. Int J Eng Appl Sci Technol 04(12):310–316. https://doi.org/10.33564/ijeast.2020.v04i12.054
140. Szandała T (2021) Review and comparison of commonly used activation functions for deep neural networks. Bio-inspired neurocomputing. Springer, Singapore, pp 203–224
141. Bai Y, Zhang H, Hao Y (2009) The performance of the backpropagation algorithm with varying slope of the activation function. Chaos Solitons Fractals 40(1):69–77. https://doi.org/10.1016/j.chaos.2007.07.033
142. Kayri M (2016) Predictive abilities of Bayesian regularization and Levenberg–Marquardt algorithms in artificial neural networks: a comparative empirical study on social data. Math Comput Appl. https://doi.org/10.3390/mca21020020
143. Roy RB et al (2021) A comparative performance analysis of ANN algorithms for MPPT energy harvesting in solar PV system. IEEE Access 9:102137–102152. https://doi.org/10.1109/ACCESS.2021.3096864
144. Namasudra S, Dhamodharavadhani S, Rathipriya R (2023) Nonlinear neural network based forecasting model for predicting COVID-19 cases. Neural Process Lett 55(1):171–191. https://doi.org/10.1007/s11063-021-10495-w
145. Smith JS, Member S, Wu B, Member S, Wilamowski BM, Fellow L (2019) Neural network training with Levenberg–Marquardt and adaptable weight compression. IEEE Trans Neural Netw Learning Syst 30(2):580–587
146. Guanabara E, Ltda K, Guanabara E, Ltda K (2022) Artificial intelligence and machine learning for EDGE computing. Elsevier. https://doi.org/10.1016/C2020-0-01569-0
147. A. Ranganathan, "The Levenberg-Marquardt Algorithm 3 LM as a blend of Gradient descent and Gauss-Newton itera," *Internet httpexcelsior cs ucsb educoursescs290ipdfL MA pdf*, 142.June, pp. 1–5, 2004, [Online]. Available: http://twiki.cis.rit.edu/twiki/pub/Main/AdvancedDipTeamB/the-levenberg-marquardt-algorithm.pdf
148. Burden F, Winkler D (2008) Bayesian regularization of neural networks. Angew Chem Int Ed 6(11):951–952. https://doi.org/10.1007/978-1-60327-101-1_3
149. Imam A, Salami BA, Oyehan TA (2021) Predicting the compressive strength of a quaternary blend concrete using Bayesian regularized neural network. J Struct Integr Maint 6(4):237–246. https://doi.org/10.1080/24705314.2021.1892572
150. Yang B et al (2021) Parameter extraction of PEMFC via Bayesian regularization neural network based meta-heuristic algorithms. Energy 228:120592. https://doi.org/10.1016/j.energy.2021.120592
151. Liu CS, Atluri SN (2011) An iterative method using an optimal descent vector, for solving an Ill-conditioned system Bx = b, better and faster than the conjugate gradient method. C Comput Model Eng Sci 80(3–4):275–298
152. Møller MF (1993) A scaled conjugate gradient algorithm for fast supervised learning. Neural Netw 6(4):525–533. https://doi.org/10.1016/S0893-6080(05)80056-5
153. Khan TA, Alam M, Shahid Z (2019) Comparative performance analysis of Levenberg–Marquardt, Bayesian regularization and scaled conjugate gradient for the prediction of flash floods. J Inform Commun Technol Robot Appl 1:52–58
154. Selvamuthu D, Kumar V, Mishra A (2019) Indian stock market prediction using artificial neural networks on tick data. Financ Innov. https://doi.org/10.1186/s40854-019-0131-7
155. Amalanathan AJ, Vasa NJ, Harid N, Griffiths H, Sarathi R (2021) Classification of thermal ageing impact of ester fluid-impregnated pressboard material adopting LIBS. High Volt 6(4):655–664. https://doi.org/10.1049/hve2.12092
156. Amizhtan SK et al (2022) Experimental study and ann analysis of rheological behavior of mineral oil-based SiO_2 nanofluids. IEEE Trans Dielectr Electr Insul 29(3):956–964. https://doi.org/10.1109/TDEI.2022.3173514
157. Vu HL, Bolingbroke D, Ng KTW, Fallah B (2019) Assessment of waste characteristics and their impact on GIS vehicle collection route optimization using ANN waste forecasts. Waste Manag 88(2019):118–130. https://doi.org/10.1016/j.wasman.2019.03.037
158. Banerjee P, Sau S, Das P, Mukhopadhayay A (2015) Optimization and modelling of synthetic azo dye wastewater treatment using graphene oxide nanoplatelets: characterization toxicity

evaluation and optimization using artificial neural network. Ecotoxicol Environ Saf 119:47–57. https://doi.org/10.1016/j.ecoenv.2015.04.022
159. Ghosal S, Chaki S (2010) Estimation and optimization of depth of penetration in hybrid CO_2 LASER-MIG welding using ANN-optimization hybrid model. Int J Adv Manuf Technol 47(9–12):1149–1157. https://doi.org/10.1007/s00170-009-2234-1
160. Ranade R, Alqahtani S, Farooq A, Echekki T (2019) An ANN based hybrid chemistry framework for complex fuels. Fuel 241:625–636. https://doi.org/10.1016/j.fuel.2018.12.082
161. Tian W, Liao Z, Zhang J (2017) An optimization of artificial neural network model for predicting chlorophyll dynamics. Ecol Modell 364:42–52. https://doi.org/10.1016/j.ecolmodel.2017.09.013
162. Garg A, Jain S (2020) Process parameter optimization of biodiesel production from algal oil by response surface methodology and artificial neural networks. Fuel 277:118254. https://doi.org/10.1016/j.fuel.2020.118254
163. Ortiz F, Simpson JR, Pignatiello JJ, Heredia-Langner A (2004) A genetic algorithm approach to multiple-response optimization. J Qual Technol 36(4):432–450. https://doi.org/10.1080/00224065.2004.11980289
164. Álvarez MJ, Ilzarbe L, Viles E, Tanco M (2009) The use of genetic algorithms in response surface methodology. Qual Technol Quant Manag 6(3):295–307. https://doi.org/10.1080/16843703.2009.11673201
165. Khoo LP, Chen CH (2001) Integration of response surface methodology with genetic algorithms. Int J Adv Manuf Technol 18(7):483–489. https://doi.org/10.1007/s0017010180483
166. Mustefa Beyan S, Venkatesa Prabhu S, Mumecha TK, Gemeda MT (2021) Production of alkaline proteases using *Aspergillus* sp. isolated from Injera: RSM-GA based process optimization and enzyme kinetics aspect. Curr Microbiol 78(5):1823–1834. https://doi.org/10.1007/s00284-021-02446-4
167. Sabry I, El-Zathry NE, Gadallah N, Abdel Ghafaar M (2022) Implementation of hybrid RSM-GA optimization techniques in underwater friction stir welding. J Phys Conf Ser. https://doi.org/10.1088/1742-6596/2299/1/012014
168. Hasanien HM, Muyeen SM (2013) A Taguchi approach for optimum design of proportional-integral controllers in cascaded control scheme. IEEE Trans Power Syst 28(2):1636–1644. https://doi.org/10.1109/TPWRS.2012.2224385
169. Ajala SO, Alexander ML (2020) Multi-objective optimization studies of microalgae dewatering by utilizing bio-based alkali: a case study of response surface methodology (RSM) and genetic algorithm (GA). SN Appl Sci 2(3):1–20. https://doi.org/10.1007/s42452-020-2097-5
170. Obi CC, Nwabanne JT, Igwegbe CA, Ohale PE, Okpala COR (2022) Multi-characteristic optimization and modeling analysis of electrocoagulation treatment of abattoir wastewater using iron electrode pairs. J Water Process Eng 49:103136. https://doi.org/10.1016/j.jwpe.2022.103136
171. Rhee JH, Il Kim S, Lee KM, Kim MK, Lim YM (2021) Optimization of position and number of hotspot detectors using artificial neural network and genetic algorithm to estimate material levels inside a silo. Sensors. https://doi.org/10.3390/s21134427
172. Chen Z, Lin X, Xiong C, Chen N (2020) Modeling the relationship of precipitation and water level using grid precipitation products with a neural network model. Remote Sens. https://doi.org/10.3390/rs12071096
173. Chiroma H et al (2017) Neural networks optimization through genetic algorithm searches: a review. Appl Math Inf Sci 11(6):1543–1564. https://doi.org/10.18576/amis/110602
174. Ferentinos KP (2005) Biological engineering applications of feedforward neural networks designed and parameterized by genetic algorithms. Neural Netw 18(7):934–950. https://doi.org/10.1016/j.neunet.2005.03.010
175. Ahmadi MA, Shadizadeh SR (2013) Retracted article: intelligent approach for prediction of minimum miscible pressure by evolving genetic algorithm and neural network. Neural Comput Appl 23(2):569. https://doi.org/10.1007/s00521-012-0984-4
176. Oreski S, Oreski D, Oreski G (2012) Hybrid system with genetic algorithm and artificial neural networks and its application to retail credit risk assessment. Expert Syst Appl 39(16):12605–12617. https://doi.org/10.1016/j.eswa.2012.05.023
177. Dong P, Liao X, Chen Z, Chu H (2019) An improved method for predicting CO_2 minimum miscibility pressure based on artificial neural network. Adv Geo-Energy Res 3(4):355–364. https://doi.org/10.26804/ager.2019.04.02
178. Ghodousian A, Babalhavaeji A (2018) An efficient genetic algorithm for solving nonlinear optimization problems defined with fuzzy relational equations and max-Lukasiewicz composition. Appl Soft Comput J 69:475–492. https://doi.org/10.1016/j.asoc.2018.04.029
179. Bahrami S, Doulati Ardejani F, Baafi E (2016) Application of artificial neural network coupled with genetic algorithm and simulated annealing to solve groundwater inflow problem to an advancing open pit mine. J Hydrol 536:471–484. https://doi.org/10.1016/j.jhydrol.2016.03.002
180. Taghavi M, Gharehghani A, Nejad FB, Mirsalim M (2019) Developing a model to predict the start of combustion in HCCI engine using ANN-GA approach. Energy Convers Manag 195:57–69. https://doi.org/10.1016/j.enconman.2019.05.015
181. Suresh MVJJ, Reddy KS, Kolar AK (2011) ANN-GA based optimization of a high ash coal-fired supercritical power plant. Appl Energy 88(12):4867–4873. https://doi.org/10.1016/j.apenergy.2011.06.029
182. Dwiputranto TH, Setiawan NA, Adji TB (2021) "DGA-based early transformer fault detection using rough set theory classifier. Int Conf Adv Mechatron Intell Manuf Ind Autom. https://doi.org/10.1109/ICAMIMIA54022.2021.9807816
183. Bülbül MA, Harirchian E, Işık MF, Aghakouchaki Hosseini SE, Işık E (2022) A hybrid ANN-GA model for an automated rapid vulnerability assessment of existing RC buildings. Appl Sci. https://doi.org/10.3390/app12105138
184. Smaali A et al (2022) Degradation of azithromycin from aqueous solution using chlorine-ferrous-oxidation: ANN-GA modeling and Daphnia magna biotoxicity test assessment. Environ Res 214(P3):114026. https://doi.org/10.1016/j.envres.2022.114026
185. Karimi H, Yousefi F (2012) Application of artificial neural network-genetic algorithm (ANN-GA) to correlation of density in nanofluids. Fluid Phase Equilib 336:79–83. https://doi.org/10.1016/j.fluid.2012.08.019

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.