Results in Engineering 16 (2022) 100637

Contents lists available at ScienceDirect

Results in Engineering

journal homepage: www.sciencedirect.com/journal/results-in-engineering

An innovative synthesis of optmization techniques (FDIRE-GSK) for generation electrical renewable energy from natural resources

Ghada S. Mohammed [a], Samaher Al-Janabi [b,*]

[a] Department of Computer Science, Collage of Science/ University of Mustansiriyah, Iraq
[b] Department of Computer Science, Faculty of Science for Women (SCIW), University of Babylon, Iraq

ARTICLE INFO

Keywords:
Intelligent data analysis
Renewable energy
Solar energy
Optmization techniques
GSK

ABSTRACT

Based on the principle that the upgrading of any nation begins by raising the level of performance of its institutions that serve the community, including the Ministry of Electricity and given the development in the field of technology, and the growing need to generate electric power, so it is found that the world has tended in recent years to generate that energy from environmentally friendly sources that do not cause CO2 emissions during generation such as solar, wind, water and other sources. Finding a solution to the problem of generating electricity energy from environmentally friendly energy is a very difficult and important issue because it is one of the basic problems in our Arab societies. The main purpose is to preserve the environment from the pollution resulting from the process of generating electricity as well as reducing the material costs in its production. This work deals with two main points: how to present the optimal feasibility study for the process of generating electricity from solar; with generated an energy at the lowest cost and the highest efficiency. The paper presens find different interval renewable energy – Gain Shraking (FDIRE-GSK) model consists of five basic stages: The first stage is the process of collecting and preparing the data to make it suitable form for the decision-making stage and included several steps, including the processing of missing values and conducting normalization work, the second stage involved develop optimization algorithms called (Gain Shraking Knowledge optmization Algorithm (GSK). That tool select after deep analysis achieve on optimization algorithms include PSOA, BOA, WOA, EHOA, ABOA and GSK; this analysis focus on determined for points: the main programming steps, main parameters, advantages, and disadvantages for each algorithm. . The GSK used to find the different intervals and remove the duplication, because GSK as algorithm having many advantages and characteristics; (a) it depends on a limited number of parameters, (b) high-accuracy performance, (c)its ability to operate with problems whose data is a collection in real-time and (d) over come the proplem of falling in local minimum; gaining and sharing knowledge base algorithm is also considered one of the optimization algorithms that (e) depends on intelligent behavior, where this algorithm distinguishes from others in that (f) it finds the important information (Knowledge or best solutions), but it only shares the most important information(optimal solutions). The FDIRE-GSK-appear as pragmatic intelligent data analysis model for reduce the computation and time of handle real big data.

1. Introduction

As a result of the development in the world of technology and information, the digital revolution in different fields, and the use of sensors, IoT devices on a daily basis to obtain and transfer data, which led to the presence of a huge amount of data that coming from different sources. Dealing with this data is for long periods, especially by programmers, who suffer from high duplication of this data and the missing of some values during the process of acquiring the data, so the processing and computations time will take a long time. Therefore, in this research, we divided the data into intervals and took the data that is within the different intervals only to significantly reduce the time of computations.

IoT considers one of the most widespread modern technologies to date, that achieve their purpose with the help of data, where the IoT devices collect data that makes business more productive and effective. The use of the Internet of things is accompanied by many challenges related to IoT data collection and management, these challenges include high sensitivity of the collected data, the lack of privacy standards, and end-to-end security solutions. Where many of the collected and

* Corresponding author.
E-mail address: samaher@itnet.uobabylon.edu.iq (S. Al-Janabi).

https://doi.org/10.1016/j.rineng.2022.100637
Received 21 June 2022; Received in revised form 9 September 2022; Accepted 11 September 2022
Available online 24 September 2022

processed information by IoT devices may be protected under various data privacy laws so must secure this protected data against attack, encryption it, manage access to it [51]. Also, there are other challenges that are represented by scalability and incorporation of renewable energy devices in the network where the huge number of devices require simultaneous connectivity to collect as much information as possible. So there are a massive amount of data produced when these online objects are synchronized with the help of sensors to support a wide range of applications [56,57]. due to the combination of large stream volume and complexity, these data suffer from duplication and the probability of missing some values, so one of the top priority challenges that related to data is managed these data and deal within it timely manner, in this research, We will find a solution to deal with this challenge through the use of several concepts that will be mentioned in the following sections.

Intelligent Data Analysis (IDA) is one of the basic and influential tools in the process of decision making due to its importance in identifying new visions and ideas, it combines different strategies and techniques to collect data from multiple sources and use it to discover knowledge and interpret it to be accurate and Internet of things (IoT) has emerged as enormous technology which finds out new different fields that have a significant impact on our life. The main objective of the IoT is to provide an environment that enables objects (such as sensors) and living beings to coordinate together to collect and share information [1, 2].

Optimization is the process of acquiring the optimal result under specific situations to minimize or maximize the objective function value of desired application [3,4]. Also, optimization refers to the process of using the resources in the best possible way, Optimization algorithms follow an appropriate representation of the problem in such a mathematical model of the optimal design to problem-based on its main components which are decision variables, constraints, and objective function. The objective optimization functions are divided into two types (single/multi) objective functions. In single-objective optimization, there is just one search space. The decision to reject or accept the solution is made based on the objective function value, While multi-objective optimization deals with finding the values of the optimal solution of (m) objective functions simultaneously [5].

Renewable Energy sources are environmental energy (alternative energy) that reduces the harmful impacts on the environment; This concept of energy is linked to the energy that is obtained from natural sources that produce enormous amounts of energy and is capable of regenerating naturally [6]; the resources of friendly environment energy are driven by the wind, hydropower, ocean waves, biomass from photosynthesis, and direct solar energy; friendly environment energy has many advantages [11,12] where it considers non-polluting, sustainable, (one-time installation), economic, ubiquitous, safe and it offers a wide variety of options and also there are drawbacks to some of the sources (Maybe more costly or its effects by some environmental influences) [7].

The remained of this paper is organized as follows: In Section 2 presents the related works, then Section 3 introduce the main concepts that are used in this work, section 4 illustrated some of the optimization techniques, Finally, in the last section we concluded the work by summarizing the main remarks and results.

2. Literature review

There are many studies that suggested many optimization algorithms to improve the methods of dealing with the large torrent of data that comes from many different sources, especially real-time data, which appears to be the same and contains a high rate of repetition and requires a long time of processing and complexity of calculations when dealing with it.

Zhao et al. [8] proposed a method for trade-off multi-objective (practical swarm optimization (MOPSO) algorithm and Techniques For Order Of Preference By Similarity To Ideal Solution (TOPSIS)) that used to achieve a strategy for energy management in system optimal configuration; also examining the strategy on the real-world case; The results show that the TPC/COE/EC set in (grid-connected, off-grid scheme) each one is optimal in different configurations. The method evaluate based on different perspectives (energy, economic, and environmental). Our work develop another optimization algorithm called (GSK) and used different dataset based work on the same types of problem is energy generation and management.

Zanabi and Samaher, (2022) suggested an integration model called "Zero to Max Energy Predictor Model based on Deep Embedded Intelligence Techniques (ZME_ DEI)" to determine the main rules for each unit that are effective in generating the maximum electrical energy based on the nature of each dataset. The ZME- DEI model creates concerning knowledge constraints and adopts gradient boosting techniques through replace the kernel of GBM represent DT by new kernel based on multi-parmater functions. The result of this model evaluation based on three measures (i.e., coefficient of determination (R2), Root Mean Square Error (RMSE), and Mean Error (ME)).). Our work different on that article by it focuess on analysis the optimization algorithm and develop GSK then adaptive it split dataset into multi interval then remove the duplication intervals while similar with it on handle the same problem.

Suman et al. [9] present an optimal techno-economic model of hybrid RES for three specific locations in Bihar state/Indian. By employing a combination of Particle Swarm Optimization and Grey Wolf Optimiser (PSO-GWO) to demonstrate a hybrid system to solve the optimal planning problem (minimized the Cost Of Electricity (COE), Deficiency of Power Supply Probability (DPSP) , higher renewable factor (RF) at the same time). In comparing with numerous other algorithms the performance of the algorithm is evaluated using,a lso using sensitivity analysis to evaluate the COE; The result shows that the best location out three is Gorigama due to its having good potential of solar and wind energy. Our work different on that article by it focuess on analysis the optimization algorithm and develop GSK then adaptive it to split dataset into multi interval then remove the duplication intervals while not take the location of solar plant into account.

Kharrich et al. [10] evaluated the project feasibility of the design of a microgrid system consisting of WT, PV, diesel, and battery; based on comparing three multi-objective optimization algorithms (Multiple Objective Particle Swarm Optimization(MOPSO), Pareto Envelope-Based Selection Algorithm(PESA II), and Strength Pareto Evolutionary Algorithm(SPEA2); with considered the multi-objective functions (Net Present Cost(NPC), Penalty Cost Of Emission, and the quantity that released into atmosphere of the CO_2))that are depended on the Six Sigma approach (set of techniques that uses probabilistic and statistical approach); and using the sensitivity analysis; The result indicates the better algorithm is SPEA2. Our work different on that article by the optimization algorithm that used and dataset; in addition; it used different evaluation measuers.

Al-Janabi et al. [11,12], suggest new methodology called (MORE-E) to generated max energy from the wind energy through compilation between Neural network and multi objectives optimization function; this model consists of five basic phases: in a first phase collecting and preparing the data, so to make it in format suitable for the decision-making stage, this phase split into multi-steps (i.e., handle missing values and normalization dataset), and the second phase focuses on building constraints for each dataset and develops one of the optimization algorithms called cuckoo based on horizontal combination and multi-objective optimization used in third phase to generate the energy. Another model is developed as multi-layer neural network called (DCapsNet) based on linear combination and multi-objective functions used in the fourth phase to generate the energy. Final phase is related to evaluation of both models (DCOM and DCapsNet) to determine the best. The MORE-G is characterized by addressing one of the real problems, saving on material costs (i.e., reducing the need for manpower and reducing dependence on other countries in importing electric power).

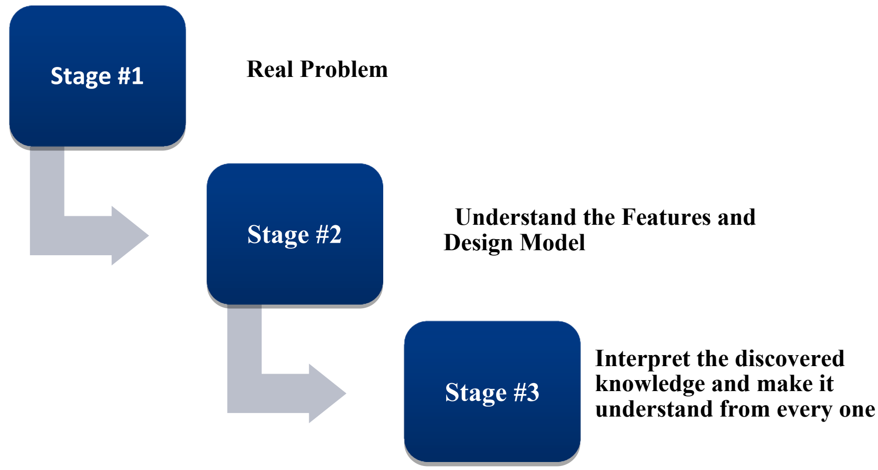

Fig. 1. Intelligent data analyses stages.

Our work different on that article through used another types of natural enery source is solar rather than wind. In addition it used GSK while that atrial used coucko optimization algorithm.

Haidar et al. [13] evaluate the performance of numerous hybrid microgrid configurations by proposing an optimal strategy and applying the hybrid system in the Long San Village/Sarawak/Malaysia. The system represents a mathematical model based on stochastic and deterministic optimization techniques (Multi-Objective Particle Swarm Optimization (MOPSO)) to support a maximum load demand at a minimum possible cost under variable weather conditions. The system behavior is evaluated by measuring the reliability and voltage security (operational analysis, techno-economic analysis, and comparative analysis of the environmental impact on the system). The targets of the system indicate by the results. Our work different on that article through dataset, techniques and evaluation measuers used but simlare with it by solve the same problem.

Kaabeche, A., & Bakelli, Y [14]. have established rules and tools for energy management optimization for the region located in the southwest of Algeria, the sizing of a separate energy production system (wind and solar) using an electrochemical storage device; techno-economic optimization norms was used to minimization the cost of energy production, by employed four optimization algorithm, also the surplus account of produced energy to maximize the efficiency of the system. The performance of the system in future is predicted, carried different sensitivity analysis and considered the effect of lifetime, depth of discharge(DOD), and the relative cost of battery technology based on the Unit Electricity Cost (UEC), JAYA algorithm produce output towards the optimal solutions more than other algorithms. Our work used another types of natural enery source is solar plant and weather rather than wind and solar. In addition it used GSK while that article used another optimization algorithm.

3. Main concepts

3.1. Intelligent data analysis (IDA)

IDA is an area of new science that is concerned with the process of analyzing the data science effectively to discover meaningful information to make a better decision for the real problem that has a specified definition and put a result in a form suitable to the user. Intelligent data analysis consists of the following major steps [1,5,51]:

- Data collection for a real problem(acquiring data, and transforming it into a suitable format for analysis)
- Preparation and Data Exploration to design a model to solve the problem based on examination specific characteristics
- Data Analysis by machine learning algorithms and the algorithms of deep learning and analysis of the results then convert it to understandable form.

The process of intelligent data analysis begins with defining the problem, determining its data, then defining and using techniques such as artificial intelligence techniques, pattern recognition, and statistics techniques to obtain the required results, and then evaluating, interpreting, and explaining these results and their impact on the decision-making process. The IDA stages illustrated in Fig (1).

3.2. Internet of Things(IoT)

IoT is a combination of Information technology and operational technology. Where to provide the IoT functionality need to the following elements(identity for each object, sensors or actuators, Communications Devices, Compute Devices, Services IoT). The main Advantages of the IoT in the real world are improved Data collection, improved customer interaction, Reduced waste, and Technology optimization. Also, there are disadvantages of using the IoT such as Flexibility, Privacy, and Complexity [15]. IoT uses sensors and actuators to interact and control the environment to achieve the specific purpose; The use of sensors and devices generate a huge amount of data about the specific features in the environment, these data must processing and discover a useful pattern from using it to modify the environment, the discovering of data pattern can perform by intelligent data analysis [2].

3.3. Optimization techniques

The aim of optimization is to find the best solution among numerous solutions that are founded in the problem space, There are many techniques to achieve the optimization aim, these techniques are classified into traditional and advanced techniques (Nayak2020), [1,16].

- Traditional techniques: deterministic optimization algorithms consist of specific rules to shift from one solution to another, as an example of these algorithms:(Linear/Nonlinear/Integer/Dynamic/Quadratic) programming, Calculus of variation, Calculus methods.
- Advanced techniques: stochastic optimization algorithms consist of rules with probabilistic transition, as an example of these algorithms: Particle Swarm Optimization Algorithm, Bat Optimization Algorithm, Whale optimization algorithm Elephant hering Optimization Algorithm, Gaining Sharing Knowledge Base Algorithm, and many others algorithms.
- The advanced techniques have features don't found in the traditional techniques also the traditional techniques have some limitations

Table 1
Comparison among some optamization techniqes.

Algorithm	Advantages	Disadvantages	Main parmaters	Secondary Parmates
Bat (Nakamura et al., 2013). [23]. (Zhao& Gao 2014) [24], (Gu erraicheet al 2021).	- The number of its parameters is specific - its structure and concept is unpretentious. -: good robustness - good exploitation ability	- The converge speed is depressed.. - Shortage of exploration ability - need to specific strategy to transition between exploration and exploitation in appropriate time - Slow convergence. - the accuracy of performance needs to enhance	- f() - Vxi - $\mathbf{F_i}$ - B_{Xi}	- Pa - D - Max_x, Min_x - R_1,R_2, and R_3 - F_{min} ,F_{max} - L_i,$\mathbf{Ra_i}$
PSO (Zhang et al., 2015) [25] [24]	- Conceptually the algorithm is very simple. - The computations are simple and derivative-free. - The algorithm is easy to implementation - Its parameters are limited and their impact to the solutions is small in comparison with other optimization techniques. - Ensure convergence and the optimum value of the problem is easily computed in a littel time.	- In high dimension problems, its easy to fall in local minimum - Determine the initial parameters depending on the principle of try and error(randomly). - In real-time applications that need high speed, the PSO may prohibit due to its iterative nature.	- f() - V_{Xi} - BP	- Pa - D - $\mathbf{Max_x}$, $\mathbf{Min_x}$ - R_1,R_2 - C1,C2 - w
WOA (Horng et al., 2017). (Kaveh & Ghazaan 2017) [29] [30] [26]	- Solving the complex optimization problems. - Its structure is simple so it used in many fields - Good capability of balancing the exploitation and exploration stages - Good performance	- Low speed in real-time high dimension optimization problems. - Easy to fall in local optimum - Low accuracy	- f() - $\mathbf{C_C}$ - $\mathbf{A_C}$ - di	- Pa - D - $\mathbf{Max_x}$, $\mathbf{Min_x}$ - $\mathbf{X_{be}}$ - R_1,R_2,$\mathbf{RX_i}$ - AU, B, **p**
EHOA [34] [35] [31] [33].	- Excellent ability for global optimization - Simple implementation. - Little number of control parameters - Difficult to stumble in local minimum	- Slow convergence. - un ability to control the balance of expolition and exploration. - Fixed parameters (scale factore is constant) - Due to its behavior, it requires more attention to he solutions	- f() - cl - $\boldsymbol{\alpha}$ - $\boldsymbol{\beta}$	- Pa - D - $\mathbf{Max_x}$, $\mathbf{Min_x}$ - R_1,R_2 - X_{be} ,X_{mi} , X_{wo}
ABOA [36]. [39] [38,40,45]	- Tries to find a solution to early convergence or stagnancy. - A little number of control parameters	- The total performance needs to improve - Need to a mechanism to keep a tradeoff between exploitation and exploration rates.	- LP_1,LP_2 - λ - w_i, m_i	- Pa - D - Max_x, Min_x - R_1
GSK [42] [44]. [43].	- Simple to implement - Limited number of parameters. - Good convergence speed - It is a dependable (relibel)technique for real-world optimization problems. - Extract the important information and identify the most important information. - Its performance resistance to curse of dimensionality.	- The algorithm is incapable of handling and solving multi-objective constrained optimization problems. - The method cannot address issues with enormous dimensions or on a wide scale.	- f() - K_f - K_r - K, p	- Pa - D - Rand - Max_x, Min_x - X_{be} ,X_{mi} ,X_{wo}

such as it not being ideal for solving multi-objective optimization problems, not being suitable for problem that has a large number of constraints, and the problem with multi-model (Patel et al., 2019), (Nayak2020).

These techniques start with generating randomly numerous candidate solutions (individuals) that represent the population, then using the objective(fitness)function to evaluate and determine the quality of the candidate solution then improving them during the iterations of the algorithm until reaching tofully implement the algorithm then candidate best solution that has best objective function value, is provided. based on the balancing between the capabilities of exploration(globally search the numerous regions of search space) and exploitation(locally search near the obtained solution) , the optamization techniques are improve to be effective and efficient to achieve the optimal solution. Trojovsky& Dehghani [17]. The optimization methods can find the optimum solution for the problem with a(single-variable function, a multivariable function with no constraints, and a multivariable function with equality or inequality constraints) [18,19]. Table (1) show a comparison among some optamization techniqes.

The optimization method can be classified based on its dealing with the individuals (solutions) in the problem search space into:

- Programming based optimization techniques that represent the population as a decision tree and each individual (program) is judged based on its ability to solve the optimization problem [20,21] from these techniques are Genetic Programming (GP)Cartesian Genetic Programming (CGP) and others
- Intelligent based optimization techniques: techniques that are inspired by the natural behavior of swarms of animals that are characteristic by their intelligence, each individual has its intelligence ability and a combination of all individual construct a strong tools to find a solution to complex problems, also these techniques can mimics the human behavior in its physical and non-physical activites [1,51]. These techniques deals with a population as randomly generated roups of solution, each group with a leader (agent) leads it during generations iterative until finding the best global solution.
- Evolutionary-based optimization techniques: techniques that produce optimal individuals based on the process of gradual improvement and change, the better generation produce based on the natural select, mutation and crossover operators, examples of these techniques are genetic algorithm (GA), Quantum evolutionary algorithm (QEA), and Backtracking optimization algorithm (BOA) and other [22]. See Fig. 2.

In our work we using *Gaining Sharing Knowledge Based Algorithm*.

3.3.1. The Bat Optimization Algorithm(BOA)

BOA is population based, metaheuristic technique that simulates the behavior of microbats for foraging. Its probosed based on the manner that the bats search for the food(prey) and avoiding the barriers. The BOA has the capabilities to find a solution for difficult, structured

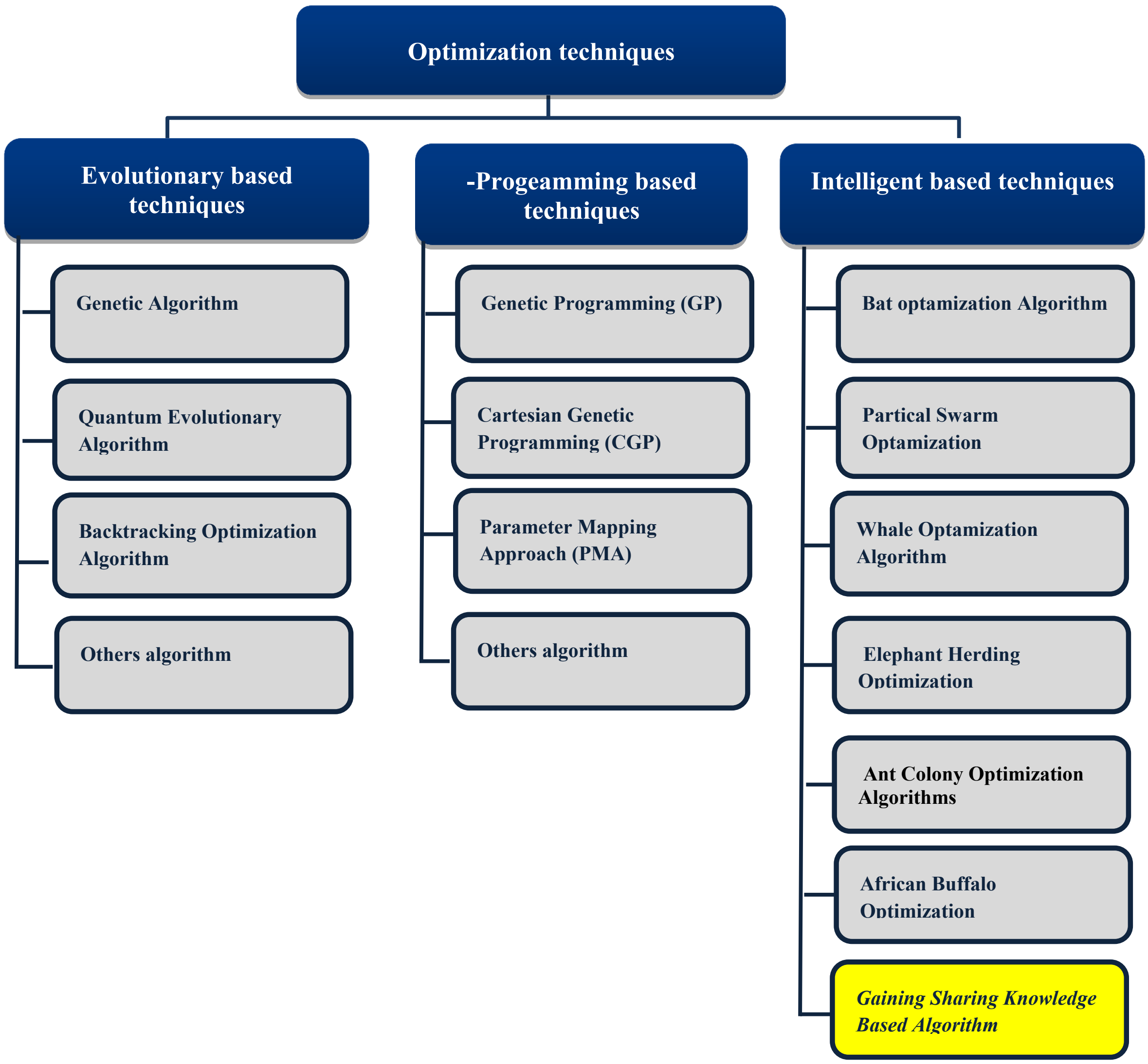

Fig. 2. Classification of Optimization techniques.

optimization problems, also has efficient ability to global search, parallel and intelligent computations because the BOA is incorporation of reinforcement search, particle swarm optimization algorithms and more parameters. The BOA has some limitations such as its difficult to use due to its using many parameters to control the global search and compute the bound of convergence, many researchers developed many versions of BOA to overcome its limitations [52]. In [53] developed BOA for continuous search space, where there are different optimization problems such as data mining, fuzzy logic, image processing etc, need to developed numerous versions BOA such as Fuzzy BOA, Binary BOA, Dicreat Binary Variant BOA, BOA with Mutations, Improved BOA, Enhanced BOA, Chaotic BOA, Compact BOA, Cloud Model BOA, and more others developed versions [23]. BOA aimed to find a solution to continuous optimization problems based on the bats ability of echolocation. The micro-bats emits high sonorous and short pulsation to sense the object's distance. From prey due to flying the bats randomly with different frequencies($\mathbf{F_i}$), velocity(VXi) at position (BXi) [23]. The echo return to the bats ears, so the bats have the ability to distinguish between the prey and the obstacles and can hunt even in a dark environment. The equations of the bat algorithm are designed and updated based on many assumptions such as the frequency(based on it to determine minimum and maximum value of the range), pulses rate based on the proximity from the prey and loudness. See algorithm (1). [24], ([54].

3.3.2. *The Particle Swarm Optimization algorithm (PSOA)*

PSOA is a well-known population based,metaheuristic technique that was inspired by the natural and biological behavior of bird swarms to solve the problems of optimization the searching for sources of food. the algorithm has many advantages such as its simplicity, easy of implementation and its ability to find a solution for different optimization problems, also the PSOA has many limitations related to its performance and early convergence to the best solution(the early convergence occure when one of the individuals find a local not global best solution and the other individuals in the population fly to it, so the algorithm is trapped with out xploiting the all search space regions), so many researcher give this algorithm more attention and made a modification to the original version it; some of them do that to overcome the algorithm limitations and some others to exploite its characteristics (Zhang et al., 2015), [25]. In (1995), Kennedy and Eberhart introduce the PSO proposed Algorithm, then numerous versions of it were developed based on defining new parameters (such as Guaranteed Convergence Particle Swarm Optimization Algorithm(GCPSOA)), hybridize it with anothers' algorithms(such as Differential Evolutionary Particle Swarm Optimization Algorithm(DEPSOA), Evolutionary Particle Swarm Optimization Algorithm(EPSOA)) or other modifications (such as Binary Particle Swarm Optimization Algorithm (BPSOA), Fully Informed Particle Swarm Optimization Algorithm (IPSOA), Stretched Particle Swarm Optimization Algorithm (SPSOA), Cooperative Particle Swarm

Optimization Algorithm (CPSOA)) ([25]. PSO techniques aim to at presenting computational intelligence by exploiting simple social behaviors, rather than plain individual cognitive capabilities to optimize the problem iteratively; it beginning with a population(swarm) of individuals(particle) each one is a candidate solution knows the global solutions in searching for optimum, until satisfied appropriate stopping criteria. The simplicity and ability of PSO make it used in many applications. But it also has a number of problems such as convergence problems and performance issues. See algorithm (2) [24], (Zhang et al., 2015), [25].

Algorithm 1
BOA

Input: *Pa, D* // Pa(population size) *individuals;;D:#features*
Output: *BP* *// BP: global best solution*

Inilization F_i , Ra_i,L_i , V_{Xi} , B_{Xi} .
While *not meeting the stop conditions* ***do*** *//stop conditions is reaching to a maximum number of iterations*
For *each i in Ba*
$F_i = F_{i-min} + R_1 * (F_{i-max} - F_{i-min})$ // adjusting F_i , updating V_{xi} , $B_{Xi,}$ *to* generate new solutions;
$V_{xi} = V_{xi-1} + (X_{be} - X_{i-1}) * F_i$
$BX_i = BX_{i-1} + V_{xi}$ *// update the location of bat*
IF $R_2 > Ra_i$
$NX_i = OX_i + L_{AL} * R_3$ *// exploiting stahe (generating the local solutions);*
End IF
IF $R_2 < Li\ and f(X_i) < f(X_{be})$
The new solution is accept
reduce Li and increase Ra_i
End IF
Arrange the solutions and select the best one.
End For
End while
End BOA

best position within the swarm (and its corresponding value in the context of the problem), Also best position of individual and its fitness value that found so far during the search process in the problem's solution space. At each iteration, represent the velocity and the position of each individual in the population, which is constantly updated to direct the iterated flights of the particles over the space of the problem possible

Where $\mathbf{f(X_i)}$ **is**a fitness function, $\mathbf{Ra_i}$:the rate of puls, Li:loudness, $\mathbf{F_{i-max}}$:maximum frequency, $\mathbf{F_{i-min}}$:minimum frequency, $\mathbf{NX_i}$: new solution, $\mathbf{OX_i}$: old solution, $\mathbf{L_{AL}}$:average of loudness of all bats during specific time,R_1,R_2, and R_3: random numbers.

Algorithm 2
PSO

Input: *Pa, D*
Output: *BP*
Inilization the population of individuals($_{Xi}$*);* p_{Xi} , V_{Xi} ,
While *not meeting the stop conditions* ***do*** *//stop conditions is reaching to a maximum number of iterations*
For *each i in Pa*
Select initial population of *particle*
Compute f($_{Xi}$*) of each particle*
IF *f(*X_i*) better than f(*X_{be}*)*
$X_{be,i} = X_i$
End IF
End For
// update the velocity of each particals
BP= X_{be}
g= $P_{Xi,}$
For *each i in Pa*
$Vxi = W*Vxi + C_1*R_1*(X_{be,i} - X_i) + C_2*R_2*(BP - X_i)$
$X_i = X_i + Vxi$
End For
End while
End PSO

Where P_{Xi},:agent position in the solution space, W:control parameter that decreased linearly,C1,C2:acceleration factors.

3.3.3. Whale optimization Algorithm(BOA)

WOA is a metaheuristics population-based technique that simulate the humpback whales' social behavior in a multi-dimension search space for hunting. This algorithm has an enormous ability to solve many optimization problems in different fields due to its properties such as simplicity, efficiency, springiness, and the power of this algorithm represent in its global exploration stage. Like many other optimization techniques, the has have some limitations such as it is easy to fall in local optimum so it will fail to perform the global search so these limitations have caused many researcher to modify the algorithm, crossbreed it with other algorithms, or enhance its parameters in ordr solve the optimization problems [26]. In (2016) WOA was developed by Mirjalili and Lewis [27] for continous search space only, where there are numerous problems of optimization need to binary search space such as data mining applications, feature selection applications and many others, so binary version of this algorithm is developed to solve these problems [28]. Another versions of WOA is developed to over come the shortcomes of low accuracy and the speed of convergence that is slow such as (IWOA) that present by Ref. [29] based on switching the probability and cofficent vectors determining ad replacing. EWOA is another version of WOA present by(Kaveh & Ghazaan 2017) where they improved the performance of algorithm by defining new parameters. MWOA multi-objective algorithm in an environment with a complex constraint is present by (Horng et al., 2017). Some other researcher hyperized WOA with another algorithm to solve constraine or in constrained problems, and even if that is considere an improvement it will increase the algorithm complexity and time complexity [30]. Many others improvement was produced and still there are more works to do esptially in real world optimization problem. WOA aimed to optimize the numerical problems in different fields, it simulate the humpback whales hunting mechanism through three principles: the first one is finding a small fish by the whale leader then shrinking encircling them and the other whales(individuals) move toward the current best prey that assumed it as a target in the search space; the second is spiraling mechanism(spiral and bubble net attack) and updating the individual locations based on the distance between current individual and optimal locations; the last principle is the prey hunting (replacing the current best location by randomly selected individual) [51]. WOA is very similar to Grey Wolf Optimization Algorithm (GWOA) exception of the hunting mechanism where in GWO algorithm the shrinking mechanism followed. See algorithm (3).

Algorithm 3
WOA

Input: *Pa, D, AU*
Output: *BP*, $f()_{be}$, X_{be}
For *each i in Pa*
    Inilization the population of agents(X_i);
    Compute the fitness $f(X_i)$ *of each individual*
End For
Select X_{i-be}
While *not meeting the stop conditions do* — *//stop conditions is reaching to a maximum number of iterations*
    For *each i in Pa* — *//computation of control cofficents*
        $A_C = 2 * AU * R_1 - AU$
        $C_C = 2 * R_2$
        IF $p < 0.5$
            IF $|A_C| < 1$
                $di = |C_C * X_{be,i} - X_i|$ — *//updating the position of whale;*
                $X_{i+1} = X_{be,i} - A_C * di$
            Else
                Selecte RX_i
                $di_2 = |C_C * RX_i - X_i|$
                $X_{i+1} = RX_i - A_C * di_2$
            End IF
        Else
            $di_1 = X_{be,i} - X_i$ — *//update the position of* X_i
            $X_{i+1} = di_1 * e^{BR_1} * \cos(2\pi R_1) + X_i$
        End IF
    End For
    Compute the fitness $f(X_i)$ *of each individual*
    Update the value of X_{be} *based on fitness value of individuals*
End while
End WOA

Where,**AU**:linearly decreased value,di: the distance between location of $\mathbf{X_{be,i}}$ and current $\mathbf{X_i}$, $\mathbf{A_C}$,$\mathbf{C_C}$:control cofficents $\mathbf{RX_i}$:random solution, B: constant number, **p**:probability;

3.3.4. The elephant herding optimization algorithm (EHOA)

EHOA is a population-based metaheuristic optimization technique that simulates the behavior of herding in the elephant groups; the algorithm has important characteristics such as it required a limited number of parameters, simple structure, and efficiency but at the same time The algorithm suffers from unbalancing between the exploration and exploitation, early convergence and easy to fall into local optimum, So to overcome these limitations the researchers made many modifications to the algorithm, the algorithm is hyperized with other algorithms to improve the initial population or add new parameters to adjust the strategy of updating the individual locations [31,32]. In (2015), wang et al. proposed EHOA, after that the original algorithm is modified ia a different method and produced different versions of the algorithm such as Improved Elephant Herding Optimization Algorithm(IMEHOA) that produce new strategy to update the position and the speed of individuals, Binary Elephant Herding Optimization Algorithm (BEHOA), three different versions of EHO were produced with alpha tuning, with culture based and with biased initialization also gradient based EHOA version was presented [33]. EHOA aimed to find a solution to different optimization tasks; algorithm is modeled into two operators the first one is the clan operator(the population of all elephants is divided into determined number of clans [58], each individual(agent) in the clan update its position according to its relationship with the leader of clan (every clan is led by the matriarch that is a female agent where its position represent the best global one in each iteration)); the second one is separating operator(simulate the properties of male elephant life that they leave their population(clan)when they adult, and he represent the worst position in each iteration. See algorithm (4) [34,35].

Algorithm 4
EHOA

Input: *Pa, D,*Max_x, Min_x

Output: *BP*

For *each i in Pa*
Inilization the population ;
Compute the fitness $f(X_i)$*of each individual*
End For
Divided the pa in to set of cl
While *not meeting the stop conditions do* *//stop conditions is reaching to a maximum number of iterations*
rank all $f(X_i)$
For *each cl in pa* *//clan operator*
For *each i in cl*
$NX_{cl,i} = X_{cl,i} + \alpha * (X_{be,cl} - X_{cl,i}) * R1$
IF $X_{cl,i} = X_{be,cl}$
$$X_{mi,cl,d} = \frac{1}{n_{cl}} * \sum_{k=1}^{n_{cl}} X_{cl,k,d}$$
$NX_{cl,i} = \beta * X_{mi,cl}$
End IF
End For
End For
For *each cl in pa* *// separating operator*
$X_{wo,cl} = Min_x + (Max_x - Min_x + 1) * R_2$
End For
Evaluate the individuals based on its position
Ene While
End EHOA

Where cl: # clans,$NX_{cl,i}$: new individuals ,α: scale factor, n: # elephant in the clan, β: random factor.

3.3.5. The african buffalo optimization Algorithm(ABOA)

ABOA is an efficient, effective metaheuristics population based technique inspired by african buffaloes alarm and alert calls during their activities of foraging and defending. ABOA used in different applications such as data mining, classifications, and many others fields. This algorithm try to solve the problem of early convergence, and if the best solution is not found after aspesific number of iteration then reinilaization the population to improve the performance and adequate the exploration stage also the ABOA over come the low speed problem by using few parameters, all that considered as properties to this algorithm [36]. Also it have some limitations such as difficicult to deal with real world proplem especially in the problem with large size and multiy constraints [37]. In (2015), Odili et al. proposed ABOA. Then this algorithm is modified by many researchers to overcome its limitations or to achieve spesific goals. The intelligent behaivor of ABOA is expliot to produce an optimal decision tree that is globally optimized by African Buffalo Optimization Algorithm Decision Tree (ABODT) [38], ABOA is hyperized with (GA) to find a solution to numerical and the graph based problems [39]. ABOA aimed to find a solution of early convergence in other optimization techniques; it has an exceptional ability for problem space exploitation(indicated by "maaa" signal)and exploration("waaa" indicate the new search); ABO algorithm updating regularly the individual's locations based on the relation of the best specific previous location of buffalo in the herd to the current location of the best buffalo; ABO has a good ability to solve numerous types of the optimization problem and provide the optimal solution faster than many other well-known algorithms such as (PSOA,GA, Ant Colony optimization algorithms(ACO)). See algorithm (5) [40,41].

Algorithm 5
ABOA

Input: *Pa, D,* LP_1 *,* LP_2
Output: *BP*

For *each i in Pa*
  Initialization the population
  Compute the fitness $f(X_i)$*of each individual*
End For
While *not meeting the stop conditions do* *//stop conditions is reaching to a maximum number of iterations*
  For *each i in Pa* *// update the individuals exploitation*
    $m_{i+1} = m_i + LP_1 * R_1\left(f\left(X_{be,i}\right) - w_i\right) + LP_2 * R_1\left(X_{be,i} - w_i\right)$
    $m_{i+1} = \frac{w_i + m_i}{\lambda}$ *// update the individuals loctions*
    IF X_{be} *updating*
      Go to 13
    Else
      Go to 2
    End IF
  End For
End While
End EHOA

Where $\mathbf{LP_1}$,$\mathbf{LP_2}$:learning factors $\mathbf{w_i}$, $\mathbf{m_i}$: moves of exploration, exploitation, λ :time unit.

3.3.6. Gaining Sharing Knowledge Based Algorithm (GSK)

GSK is a metahuristic technique that based on population and a nature inspired and it considered the behavior of human life spin. this algorithm characterized by its robustness, the overall performance is stable(roubest) even the dimension increase, good convergence speed, the high quality problem solution even with high dimension problems, complex problems and real time optamization problems [42], like other optamization algorithm, the GSKA have some of limitations such as it work in continuous search space only and can not operate in discreat search space directly, also if the population is not diversity the algorithm is fall in early convegrous, so the researcher exploite this limitation to present improved versions of GSKA [43]. In (2020), Mohamed et al. proposed GSK algorithm. then developed by adaptive the important parameter(knowledge factor($\mathbf{K_f}$) ,**knowledge ratio**($\mathbf{K_r}$)) that control junior and senior stages for gaining and sharing knowledge to produce adaptive Gaining Sharing Knowledge Based Algorithm(IGSKA) [42]; Also,a novel binary version of Gaining Sharing Knowledge Based Algorithm (NBGSKA) was presented to find a solution to binary search space problems then this algorithm was developed to enhance the performance and increase the accuracy and avoid falling the algorithm in early convergence by producing the Population Reduction Novel Binary GSKA (PR-NBGSKA) [44]. This algorithm aimed to solve constrained and unconstrained optimization problems (especially real-world problems). GSK algorithm simulates the process of knowledge acquiring and sharing during the life span of human. GSK is composed of two stages, junior(beginning) and senior (expert)gaining and sharing stages, each stage required specific dimensions number determined based on the value of knowledge factor(K_f),K_f control the volume of acquired knowledge to the individual (where the number of dimensions that required for the junior stage is greater than the number that required to the senior stage), also the number of dimension decrease and increase linearly or non-linear based on the knowledge factor(K), to determine the fast of acquiring the experience (control the transferred knowledge that shred during the iterations of the algorithm) [55]. In the first stage, all individuals in the population interact with each other and each individual gains different types of knowledge from its specific environment although the individuals in this stage still with little experience, he also has the ability to classify the other individuals to best, middle, worst, so in next stage he selects the appropriate individuals with best characteristics and behaviors to share his knowledge in different fields with them. See algorithm (6) [42]

Algorithm 6
GSK

Input: *Pa,d*
Output: *BP*
Select initial population
Compute the fitness of each individual
For *each GE in* $\boldsymbol{GE^{max}}$
Compute the # dimensions of junior& senior stage
For *each i in Pa* *// junior stage*
For *each c in D*
IF $(rand) \leq (K_f)$
IF $f(x_{ic}) > f(X_r)$
$\boldsymbol{NX = X_{ic} + K_f * [(X_{i-1} - X_{i+1}) + (X_r - X_{ic})]}$
Else
$\boldsymbol{NX_{ic} = X_{ic} + K_f * [(X_{i-1} - X_{i+1}) + (X_{ic} - X_r)]}$
End IF
Else
$\boldsymbol{NX_{ic} = OX_{ic}}$
End IF
End for
End for
For *each i in Pa* *// senior stage*
For *each c in D*
IF $(rand) \leq (K_r)$
IF $\boldsymbol{f(X_{ic}) > f(X_{mi})}$
$\boldsymbol{NX_{ic} = X_{ic} + K_f * [(X_{be} - X_{wo}) + (X_{mi} - X_{ic})]}$
$\boldsymbol{NX_{ic} = X_{ic} + K_f * [(X_{be} - X_{wo}) + (X_{ic} - X_{mi})]}$
End IF
Else
$\boldsymbol{NX_{ic} = OX_{ic}}$
End IF
End for
End for
For *each i in Pa* *//*
IF $(f(NX_{ic}) \leq f(OX_{ic}))$
update each individual
End IF
IF $(f(NX_{ic}) \leq f(X_{be,G}))$
update the global best solution
End IF
End for
End for
End GSK

Where $\mathbf{X_r}$: random number,

4. Find different interivals of renewable energy-gaining sharing knowledge (FDIRE-GSK-) model

The FDIRE-GSK- model build through the following stages explain in algorithm (7): First Stage: Collect the data in real-times from multi sensors include seven related to Solar-Plant dataset and six related to Weather dataset. Second Stage: Merage both datasets into one basd on the primary keys (Plant-Id and Date-time). Third stage: Checking if that dataset containe missing values in any record then drop that records otherwise save that recoreds in cleaning dataset. Four stage: apply GSK- to split the clean dataset into multi intervals then select only the different intervaluse this meaning remove the duplication of intervals this steps will reduce the computation that need in the next stages. Final stage: using the Divied boulding index as measuer to Verafication from obtained results. Block diagram of model explain in Fig. 3. While Fig. 4. Shown the Flowchart of FDIRE-GSK- Model.

Algorithm 7
FDIRE-GSK-

Input: *SPD with WD____// SPD: solar plant dataset; WD: weather dataset*
Output: *Buffer of different intervals related to merged datasets*
For each i in Pa
  For each c in D
    Merge ((d$_{i,c}$_SPD)&((d$_{i,c}$_WD) in new dataset(SPDWD)
  End for
End for
For each i in Pa
  For each c in D
    IF there are missing values in (SPDWD)
    Drop record (i)
  End for
End for
For each i in Pa // compute the fitness function for all individuals in the population
  $f(X_i) = -20.0 * e^{\left(-0.2*\sqrt{\frac{1}{D}(\sum_{c=1}^{D} x_{i,c}^2)}\right)} - e^{\left(\frac{1}{D}*\sum_{c=1}^{D}\cos c*x_{i,c}\right)+20+e^{(1)}}$
End for
GE=0
For each GE in GE^{max}
  $D_{JU} = D*\left(1-\frac{GE}{GE^{max}}\right)^K$ //# dimensions of junior stage
  $D_{SE} = D - D_J$ **//# dimensions of senior stage**
  Arrange all $f(X_i)$ in ascending order.
  For each i in Pa
    For each c in D
      Select X_{i-1}, X_{i+1}
      Select (X_r)
      IF $(rand) \leq (K_f)$
        IF $f(X_{ic}) > f(X_r)$
          $NX_{ic} = X_{ic} + K_f * [(X_{i-1} - X_{i+1}) + (X_r - X_{ic})]$
          Else
          $NX_{ic} = X_{ic} + K_f * [(X_{i-1} - X_{i+1}) + (X_{ic} - X_r)]$
        End IF
      Else
        $NX_{ic} = OX_{ic}$
      End IF
    End for
  End for
  Arrange all $f(X_i)$ in ascending order.
  Select p
  $X_{be} = 100\,p\%$
  $X_{mi} = Pa - 2 * 100\,p\%$
  $X_{wo} = 100\,p\%$
  For each i in Pa
    For each c in D

To evaluate the fitness of each individuals(solutions) in the population we used three different functions (Bohachevsky, Salomon, Ackely), these functions used to test the optamization algorithms.

- B**ohachevsky function:** this function has many properties such as it is continuous, multimodal, separable, differentiable and non-scalable; This function need to nolrmalized data to spesific rang to operate with it, these step increae the compution so the time complexity will also increase (Jamil& Yang 2013).

$$f(X_i) = x_1^2 + 2x_1^2 - 0.3*\cos(3\pi x_1) - 0.4*\cos(3\pi x_2) + 0.7 \quad (1)$$

- **Salomon function:** continuous,multimodal, separable, differentiable and non-scalable all these properties of Salomon function; this function need the data to normalized to specific range to evaluate them, this causes to increase the complexity of time and computation [46].

$$f(X_i) = 1 - \cos\left(2\pi\sqrt{\left(\sum_{i=1}^{D} x_i^2\right)}\right) + 0.1*\sqrt{\left(\sum_{i=1}^{D} x_i^2\right)} \quad (2)$$

- **Ackley function:** this function is continuous, multimodal, separable, differentiable and non-scalable;

$$f(X_i) = -a*e^{\left(-b*\sqrt{\frac{1}{D}\left(\sum_{c=1}^{D} x_i^2\right)}\right)} - e^{\left(\frac{1}{D}*\sum_{c=1}^{D}\cos\,c*x_i\right)+a+e^{(1)}} \quad (3)$$

There are many research recommended the values of a = 20,b = 0.2, c = 2π as optimal values to find the global optimal solution (Hussain et al., 2017). Ackely function considered appropriate function to our works due to its ***properties***, where it is suitable to ***operate with real stream of data***, ***huge data***, it need to ***limited number of parameters***, provide ***high accuracy***, its ***computation is relatively simple,don't need transformation or normalization of data*** to operate on it ***so it less time***

IF $(rand) \leq (K_r)$
IF $\boldsymbol{f(x_{ic}) > f(X_{mi})}$
$\boldsymbol{NX_{ic} = X_{ic} + K_f * [(X_{be} - X_{wo}) + (X_{mi} - X_{ic})]}$

$\boldsymbol{NX_{ic} = X_{ic} + K_f * [(X_{be} - X_{wo}) + (X_{ic} - X_{mi})]}$
End IF
Else
$\boldsymbol{NX_{ic} = OX_{ic}}$
End IF
End for
End for
For *each i in Pa* // update each individuals
IF *(f(* NX_{ic}*)* $\leq$ *f(*OX_{ic}*))*
$\boldsymbol{OX_{ic} = NX_{ic}}$
$\boldsymbol{f(OX_{ic}) = f(NX_{ic})}$
End IF
IF *(f(* NX_{ic}*)* $\leq$*f(*$X_{be,G}$*))* // update the global best
$X_{be,G} = NX_{ic}$
$f(X_{be,G}) = f(NX_{ic})$
End IF
End for
End for
Call *the Davies boulding index*
End FDIRE-GSK

complexity, less computation complexity also, the function can used ***single optamiztion or multi optamization*** (throught linear combination).

In this work used Davies Bouldin Index (DBI) that know as one of the measures that are considered popular to evaluate the performance clustering (grouping) algorithms, this index is based on the ratio of intra-similarity and inter-similarity between clusters where the DB index has a positive relation with intra-similarity (reduce the distance between points of the cluster) and negative relation with inter-similarity (maximize the distance between clusters) [47], the popularity of using this index is due to its way of clustering validation where it includes internal and external validation to evaluate the clustering perfrmance [48] the smaller value of this index close to 0 (not negative) means the better (optimal) separations between clusters as shown in algorithm (8) [49].

Algorithm 8
DBIA

Input: *Buffer contain different intervals*
Output: *AS* // *the similarity avarege among the interival and most similar one to it*

//compute(intra-sim) the similarities among the point of interval

$$IS_i = \left(\frac{1}{|IN_i|}\sum_{X \in IN_i} |X - CI_i|_p^q\right)^{\frac{1}{q}} \qquad \text{//where;} \quad CI_i = \frac{\sum_{X \in IN_i} X}{NX_i}$$

//compute the sepation measures (distance) between of interivals(clusters)

$DI_{i,j} = \|CI_i - CI_j\|_p$

//compute the most similarity (inter-sim)among the intervals

$$inter - sim(ISS_{i,j}) = MAX_{i \neq j}\left(\frac{IS_i + IS_j}{DI_{i,j}}\right)$$

//compute Davies bouldin index (DBI)

$$AS = (1/N) * \sum_{i=1}^{N} ISS_i$$

End DBI

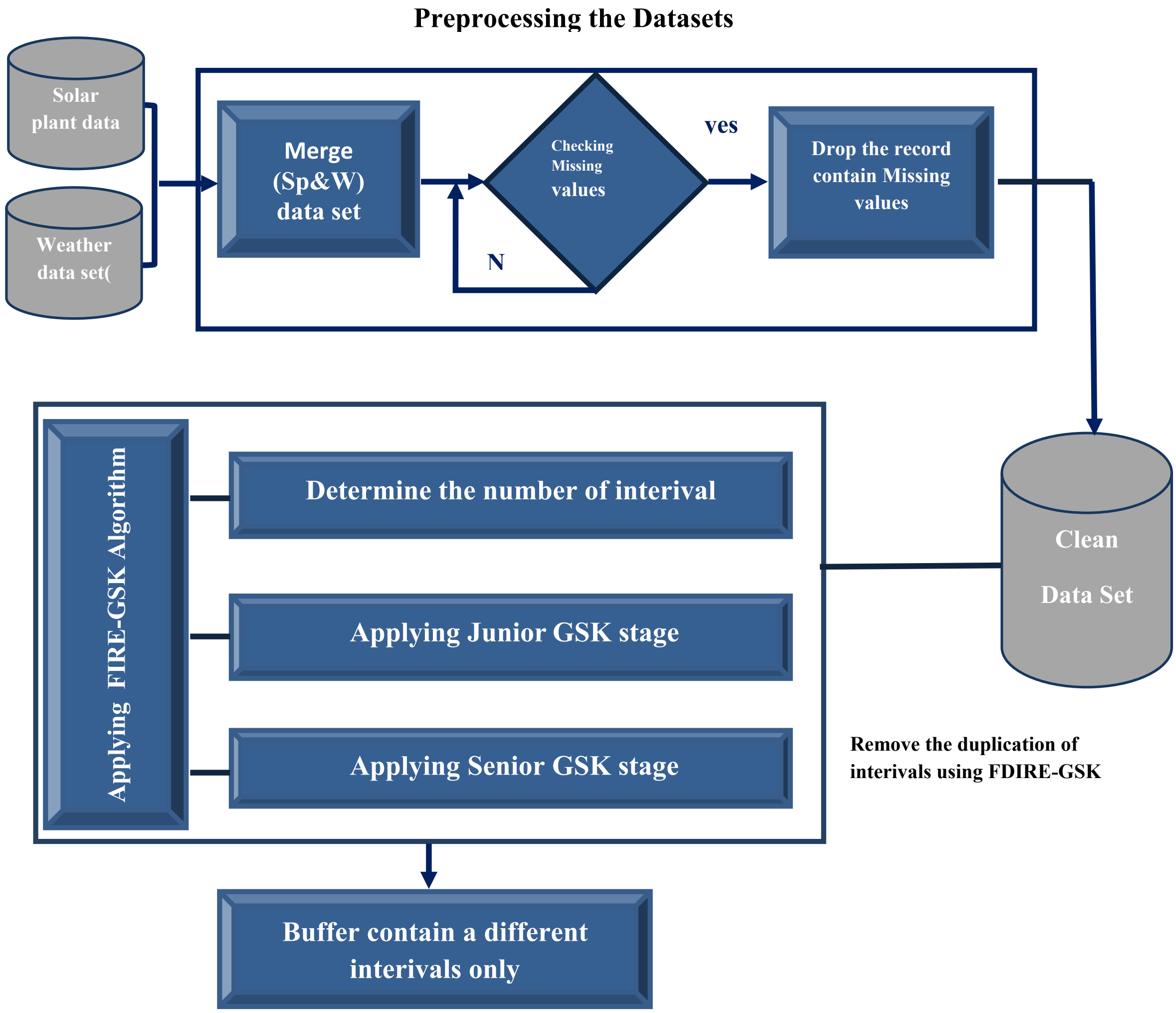

Fig. 3. Block diagram of FDIRE-GSK Model.

Where IS stand for intra-sim,IN is interval (cluster),*X*is interival point (cluster item), CI is the centroid of cluster, NX is cardinality of IN_i (number of points in the interval), DI is the distance between intervals, ISS (inter-sim) is similarities among the intervals, N is number of interivals [50].

5. Results of FDIRE-GSK-model

5.1. Collecting datasets

In this stage, the datasets collected through 5 months are used to build and test the proposed model, there are two datasets (solar and weather dataset) each one of these datasets has different features; the solar dataset consists of 68,778 samples while the weather dataset consists of 3182 samples. The solar panel dataset contains seven features [date_time, plant_id, source_key, dc_power, ac_power, daily_yield, total_yield], whereas the weather dataset includes six features [date_time, plant_id,source_key,ambient_temperature,]module_temperat-ure, irradiation). Tables 2 and 3 show the sample of weather and solar plant consequently.

5.2. PreProcessing

This stage involved many steps to handle problems related to the dataset to extract useful information and get thesis' target quickly as to show this later.

- Merging

To apply merging between two datasets based on D_T and P_ID features we need firstly adjust the D_T format to each dataset from the (day, month, year, hour, and minutes) formatting (i.e, %d-%m-%Y %H:%M) to the (year, month, day, hour, minutes, second) formatting (i.e, %Y-%m-%d %H:%M:%S) to make standard formatting in the python programming language. Drop unnecessary features and apply merging through the D_T feature.

The results of the step merging that are used to join the original datasets within one dataset along the column and rows. The main purpose of the merging step is to fill gaps in data, increase the samples size, and improve the robustness of results.

- **Checking Missing Values**

Search for the missing value is applied to drop it with its row because the dataset that dealing with contains a stream of data therefore deleting one or more rows does not affect the prediction over a long time. We must point out that the datasets dealing with does not include missing values. The main purpose of this step is to avoid training failing and make the prediction more accurate.

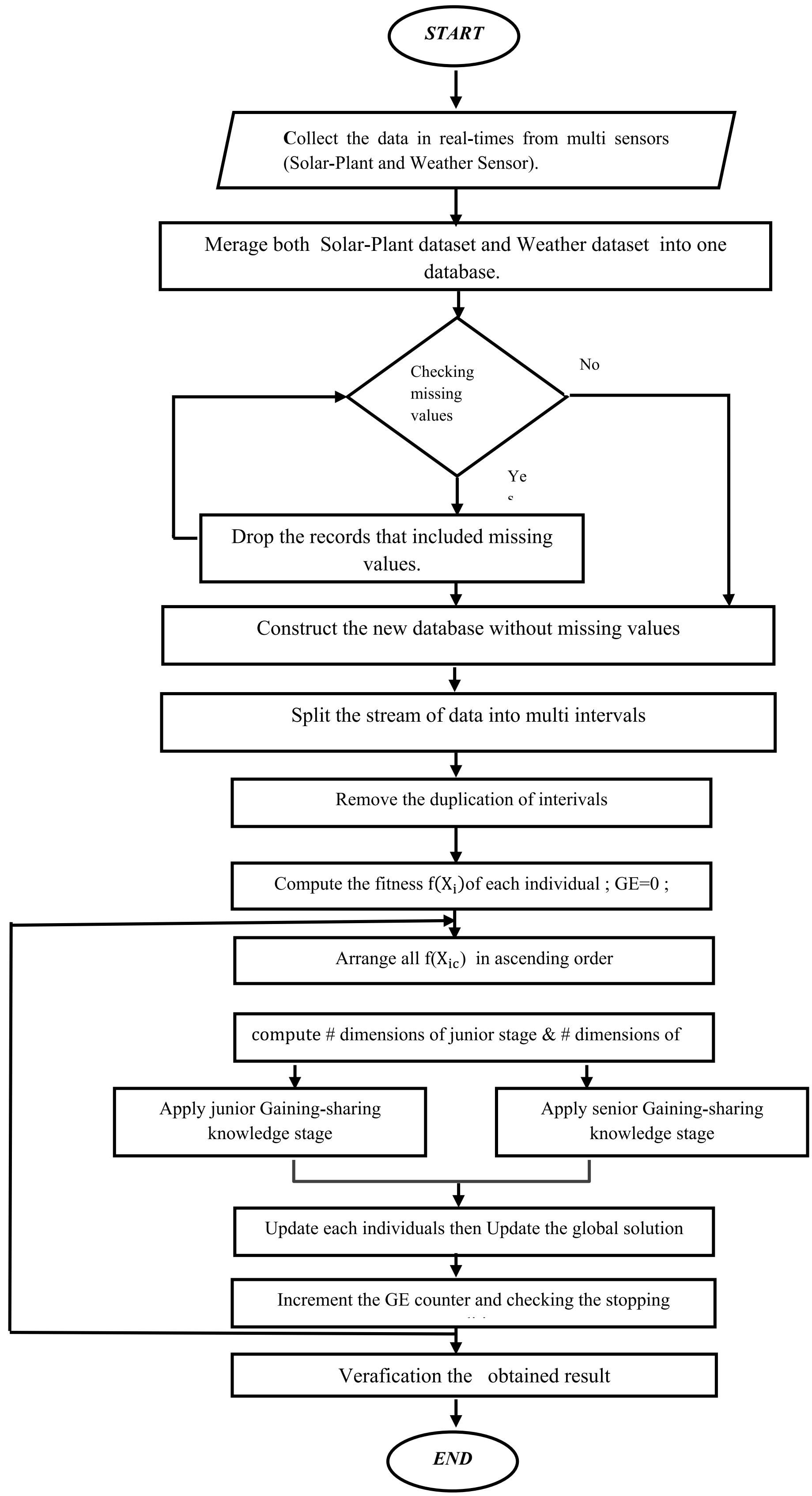

Fig. 4. Flowchart of FDIRE-GSK model.

Table 2
Sample of weather dataset.

IRR	MODULE_TEM	AMBIENT_TEM	SOU_K	P_ID	D_T
0.000862721	23.09669193	24.28921113	HmiyD2TTLFNqkNe	4135001	May 15, 2020 05:45
0.005886957	22.2067566	24.08844607	HmiyD2TTLFNqkNe	4135001	May 15, 2020 06:00
0.022281607	22.35345867	24.01163527	HmiyD2TTLFNqkNe	4135001	May 15, 2020 06:15
0.049409724	22.893282	23.97673127	HmiyD2TTLFNqkNe	4135001	May 15, 2020 06:30
0.095394454	24.44244393	24.21899	HmiyD2TTLFNqkNe	4135001	May 15, 2020 06:45
0.141940443	27.18565287	24.5373984	HmiyD2TTLFNqkNe	4135001	May 15, 2020 07:00
0.154712676	28.88847786	24.8159595	HmiyD2TTLFNqkNe	4135001	May 15, 2020 07:15
0.148799153	29.6056438	24.98878987	HmiyD2TTLFNqkNe	4135001	May 15, 2020 07:30
0.000862721	23.09669193	24.28921113	HmiyD2TTLFNqkNe	4135001	May 15, 2020 05:45
0.005886957	22.2067566	24.08844607	HmiyD2TTLFNqkNe	4135001	May 15, 2020 06:00

Table 3
Sample of solar dataset.

T_YIELD	D_YIELD	AC_P	DC_P	SOU_K	P_ID	D_T
6987759	0	0	0	3PZuoBAID5Wc2HD	4135001	May 15, 2020 00:15
7602960	0	0	0	7JYdWkrLSPkdwr4	4135001	May 15, 2020 00:15
7158964	0	0	0	McdE0feGgRqW7Ca	4135001	May 15, 2020 00:15
7206408	0	0	0	VHMLBKoKgIrUVDU	4135001	May 15, 2020 00:15
7028673	0	0	0	WRmjgnKYAwPKWDb	4135001	May 15, 2020 00:15
6522172	0	0	0	ZnxXZZZa8U1GXgE	4135001	May 15, 2020 00:15
7098099	0	0	0	ZoEaEvLYb1n2sOq	4135001	May 15, 2020 00:15
6271355	0	0	0	adLQvlD726eNBSB	4135001	May 15, 2020 00:15
6987759	0	0	0	3PZuoBAID5Wc2HD	4135001	May 15, 2020 00:15
7602960	0	0	0	7JYdWkrLSPkdwr4	4135001	May 15, 2020 00:15

Table 4
Main parmaters of GKSA

Parameter	Value
Population size(N)	All samples with packet of windows equel 100 samples
knowledge factor(k_f)	0.5
Knowledge ratio(k)	0.9
Knowledge rate(k_r)	10
Fitness function (Ackley)	$f(X_i) = -a * e^{\left(-b*\sqrt{\frac{1}{D}\left(\sum_{c=1}^{D} x_i^2\right)}\right)} - e^{\left(\frac{1}{D} * \sum_{c=1}^{D} \cos c*x_i \right)} + a + e^{(1)}$
Bohachevsky function	$F(X_i) = x_1^2 + 2x_1^2 - 0.3 * \cos(3\pi x_1) - 0.4 * \cos(3\pi x_2) + 0.7$
Salomon function	$f(X_i) = 1 - \cos\left(2\pi\sqrt{\left(\sum_{i=1}^{D} x_i^2\right)}\right) + 0.1 * \sqrt{\left(\sum_{i=1}^{D} x_i^2\right)}$
P	0.1
A	20
B	0.2
C	2π
Generations	3000

- Apply FDIRE-SGKA

Splitting data rows into multi intervals based on apply SGK and keeping the difference intervals only. The main parmaters of GSK shown in table.

From the above found total number of intervals is 296 intervals while after remove the duplication remined number of different intervals is 50 and number of samples related to each of that in tervals explain in Table 6.

6. Conclusions

This section explains the main conclusions found through implementing the FDIRE-GSK Model. Can summarization as the following:

A. It solves one of main problem related to generate the renewable energy through building-integrated software to generate max renewable energy from friendly environment resources.
B. Find the solution of the three challenges represent in the triangular problem that has three different sides though (reduce time of implantation problem this achieve through working on different intervals only, increase the accuracy of results by remove the samples that have missing values and determined the main features impact to the DC-Power by correlation problem, finally, reduce the

Table 5
Results of FDIRE-GSK based on different fitness function.

# Gneration	Senior Dimension	Junior Dimension	Ackley function	Bohachevsky function	Salomon function
100	10.288	0.712	0.271	0.948316	0.653987
200	9.616	1.384	0.379	0.768096	0.810514
300	8.983	2.017	0.395	0.691244	0.739011
400	8.387	2.613	0.528	0.736384	0.071303
500	7.298	3.702	0.350	1.243624	0.094211
600	6.801	4.199	0.221	0.768727	0.507,710
700	6.334	4.666	0.237	0.750606	0.42457
800	5.896	5.104	0.374	0.596191	0.31315
900	5.484	5.516	0.410	0.031289	0.73705
1000	5.097	5.903	0.434	0.906304	0.393793
1100	4.735	6.265	0.459	0.788114	0.570421
1200	4.396	6.604	0.527	0.025965	0.017433
1300	4.077	6.923	0.669	0.152017	0.35317
1400	3.78	7.22	0.686	0.68,516	0.368308
1500	3.501	7.499	0.726	0.288448	0.519971
1600	3.241	7.759	0.767	0.097893	0.328411
1700	2.997	8.003	0.767	0.429179	0.673543
1800	2.77	8.23	0.767	0.535565	0.298519
1900	2.558	8.442	0.767	0.532015	0.589012
2000	2.361	8.639	0.767	0.305856	0.016779
2100	2.176	8.824	0.767	0.80384	0.78,274
2200	2.005	8.995	0.767	0.120083	0.70,502
2300	1.845	9.155	0.767	0.42946	0.107644
2400	1.697	9.303	0.767	0.717078	0.006981
2500	1.559	9.441	0.767	0.829829	0.492227
2600	1.432	9.568	0.781	0.789034	0.067597
2700	1.313	9.687	0.781	0.63551	0.545198
2800	1.203	9.797	0.781	0.083048	0.760483
2900	1.101	9.899	0.808	0.454223	0.369116
3000	1.007	9.993	0.931	0.889799	0.871403

Table 6
Results of FDIRE-GSK- based on different intervals.

# Interval	# Points related to each Interval	# Interval	# Points related to each Interval
Interval #1	3202	Interval #26	415
Interval #2	3067	Interval #27	414
Interval #3	3010	Interval #28	401
Interval #4	3007	Interval #29	350
Interval #5	2996	Interval #30	326
Interval #6	2996	Interval #31	324
Interval #7	2978	Interval #32	321
Interval #8	2961	Interval #33	320
Interval #9	2950	Interval #34	318
Interval #10	2939	Interval #35	307
Interval #11	2913	Interval #36	297
Interval #12	2895	Interval #37	285
Interval #13	2825	Interval #38	279
Interval #14	2806	Interval #39	262
Interval #15	2771	Interval #40	252
Interval #16	2754	Interval #41	249
Interval #17	2751	Interval #42	246
Interval #18	2699	Interval #43	239
Interval #19	2643	Interval #44	235
Interval #20	2637	Interval #45	233
Interval #21	2598	Interval #46	232
Interval #22	2523	Interval #47	193
Interval #23	648	Interval #48	190
Interval #24	476	Interval #49	188
Interval #25	417	Interval #50	150

computation by remove duplication from intervals and work on the different intervals only).

C. The pre-processing stage involves three stages; first stage involves managing both datasets, checking of missing value, then split the stream of dataset into multi-intervals and remove the duplication this stage enhance the performance of design software and reduce the computation of total system.

D. GSK is one of the pragmatic tools to work with real data, where, GSK characteristic thorny working in parallel and give high accuracy. In general; it is based on three parameters (Ackley function, Junior Phase, Senior Phase). Therefore, developing GSK get high accuracy results but on the other side, the computation is increased on other side it reduces implementation time.

Table 7
Validation of FDIRE-GSK- based on DBI, accuracy and Time.

# Interval	DBI	Accury	Time	# Interval	DBI	Accury	Time
Interval #1	0.9278	0.81024	5s	Interval #26	0.6616	0.7721	4s
Interval #2	0.8953	0.8639	2s	Interval #27	0.7355	0.8046	3.5s
Interval #3	0.9447	0.7478	2s	Interval #28	0.6139	0.8552	2.5s
Interval #4	0.5997	0.9781	2s	Interval #29	0.7737	0.9400	2.5s
Interval #5	0.82782	0.9816	2s	Interval #30	0.6306	0.9797	2.5s
Interval #6	0.8897	0.8102	2s	Interval #31	0.9278	0.9833	2.5s
Interval #7	0.7136	0.6723	2s	Interval #32	0.8953	0.9826	2.5s
Interval #8	0.6252	0.9811	2s	Interval #33	0.9447	0.9838	2.5s
Interval #9	0.989	0.9810	2s	Interval #34	0.5997	0.9822	2.5s
Interval #10	0.8713	0.9832	2s	Interval #35	0.82782	0.9836	2s
Interval #11	0.8022	0.8177	2s	Interval #36	0.7136	0.81024	2s
Interval #12	0.8374	0.82,367	2s	Interval #37	0.6252	0.9823	2s
Interval #13	0.8022	0.8233	3s	Interval #38	0.8953	0.9811	2s
Interval #14	0.8374	0.81,941	3s	Interval #39	0.9447	0.8947	2s
Interval #15	0.6616	0.9840	3s	Interval #40	0.5997	0.9767	2s
Interval #16	0.7355	0.9823	2s	Interval #41	0.6139	0.9806	2s
Interval #17	0.7529	0.8202	2s	Interval #42	0.7737	0.8177	2s
Interval #18	0.5257	0.8233	2s	Interval #43	0.8490	0.82,367	3s
Interval #19	0.7258	0.8224	2s	Interval #44	0.7849	0.8224	3s
Interval #20	0.9351	0.8947	2s	Interval #45	0.9730	0.8947	3s
Interval #21	0.8490	0.9767	2s	Interval #46	0.6758	0.9822	2s
Interval #22	0.7136	0.9806	3s	Interval #47	0.8964	0.9836	2s
Interval #23	0.6252	0.9815	3s	Interval #48	09822	0.9781	2s
Interval #24	0.6139	0.81024	3s	Interval #49	0.9111	0.9816	2s
Interval #25	0.7737	0.9823	3s	Interval #50	0.9429	0.8102	2s

While the main limitation of the work can summarization as follow; it not take the location of solar plant into account, some time the solar plant attack, except for the number of parameters that affect the amount of energy produced from them, such as dust, birds build nests on them, throwing waste on the surfaces of solar plants, and other natural parameters are not taken into account in this article.

Credit author statement

All authors contributed to the study's conception and design. Design the system achieved by [Samaher Al-Janabi]. Test and analysis were performed by [Samaher Al-Janabi and Ghada S. Mohammad]. The first draft of the manuscript was written by [Samaher Al-Janabi] and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

[1] Samaher Al-Janabi, Ayad Alkaim, A novel optimization algorithm (Lion-AYAD) to find optimal DNA protein synthesis, Egypt. Inform. J. (2022), https://doi.org/10.1016/j.eij.2022.01.004.
[2] M. Alam, A.K. Skahil, S. Khan, Internet of Things (IoT), Concepts and Applications, Cham, Switzerland, 2020.
[3] Reena Sharma, Hariprasad Kodamana, Manojkumar Ramteke, Multi-objective dynamic optimization of hybrid renewable energy systems, Chem Eng Proc. Inten 170 (2022), https://doi.org/10.1016/j.cep.2021.108663.
[4] S.S. Rao, Engineering Optimization: Theory and Practice, John Wiley & Sons, , 2019.
[5] S. Al-Janabi, A.F. Alkaim, A nifty collaborative analysis to predicting a novel tool (DRFLLS) for missing values estimation, Springer, Soft Comput 24 (1) (2020) 555–569, https://doi.org/10.1007/s00500-019-03972-x.
[6] K. Stergiou, T.E. Karakasidis, Application of deep learning and chaos theory for load forecasting in Greece, Neural Comput. Appl. 33 (2021) 16713–16731, https://doi.org/10.1007/s00521-021-06266-2.
[7] T. Baydyk, E. Kussul, D.C. Wunsch II, Intelligent Automation in Renewable Energy, Springer International Publishing, 2019, https://doi.org/10.1007/978-3-030-02236-5.
[8] P. Zhao, F. Gou, W. Xu, J. Wang, Y. Dai, Multi-objective optimization of a renewable power supply system with underwater compressed air energy storage for seawater reverse osmosis under two different operation schemes, Renew. Energy 181 (2022) 71–90, https://doi.org/10.1016/j.renene.2021.09.041. .
[9] G.K. Suman, J.M. Guerrero, O.P. Roy, Optimisation of solar/wind/bio-generator/diesel/battery based microgrids for rural areas: a PSO-GWO approach, Sustain. Cities Soc. 67 (2021), 102723, https://doi.org/10.1016/j.scs.2021.102723. .
[10] M. Kharrich, O.H. Mohammed, N. Alshammari, M. Akherraz, Multi-objective optimization and the effect of the economic factors on the design of the microgrid hybrid system, Sustain. Cities Soc. 65 (2021), 102646, https://doi.org/10.1016/j.scs.2020.102646.
[11] S. Al-Janabi, A.F. Alkaim, Z. Adel, An Innovative synthesis of deep learning techniques (DCapsNet & DCOM) for generation electrical renewable energy from wind energy, Soft Comput. 24 (14) (2020) 10943–10962, https://doi.org/10.1007/s00500-020-04905-9. .
[12] S. Al-Janabi, S. Alhashmi, Z. Adel, Design (More-G) model based on renewable energy & knowledge constraint, in: Y. Farhaoui (Ed.), Big Data and Networks Technologies. BDNT 2019. Lecture Notes in Networks and Systems, vol. 81, Springer, Cham, 2020, https://doi.org/10.1007/978-3-030-23672-4_20.
[13] A.M. Haidar, A. Fakhar, A. Helwig, Sustainable energy planning for cost minimization of autonomous hybrid microgrid using combined multi-objective optimization algorithm, Sustain. Cities Soc. 62 (2020), 102391, https://doi.org/10.1016/j.scs.2020.102391.
[14] A. Kaabeche, Y. Bakelli, Renewable hybrid system size optimization considering various electrochemical energy storage technologies, Energy Convers. Manag. 193 (2019) 162–175, https://doi.org/10.1016/j.enconman.2019.04.064. .
[15] S.L. Peng, S. Pal, L. Huang (Eds.), Principles of Internet of Things (IoT) Ecosystem: Insight Paradigm, Springer International Publishing, 2020 ,, pp. 263–276.
[16] He Yi, Su Guo, Jianxu Zhou, Jilei Ye, Jing Huang, Kun Zheng, Xinru Du, Multi-objective planning-operation co-optimization of renewable energy system with hybrid energy storages, Renew. Energy 184 (2022) 776–790, https://doi.org/10.1016/j.renene.2021.11.116.
[17] P. Trojovsky, M. Dehghani, Hybrid Leader Based Optimization: A New Stochastic Optimization Algorithm for Solving Optimization Applications, 2022.
[18] V.K. Patel, V.J. Savsani, M.A. Tawhid, Thermal System Optimization, Springer, Cham, 2019.
[19] Guangyou Zhou, Zhiwei Zhu, Sumei Luo, Location optimization of electric vehicle charging stations: based on cost model and genetic algorithm, Energy 247 (2022), https://doi.org/10.1016/j.energy.2022.123437.
[20] A. Elkasaby, A. Salah, E. Elfeky, Multiobjective optimization using genetic programming: reducing selection pressure by approximate dominance, in: ICORES, 2017, February, pp. 424–429.
[21] R.C. Póvoa, A.S. Koshiyama, D.M. Dias, P.L. Souza, B.A. Horta, Unimodal optimization using a genetic-programming-based method with periodic boundary conditions, Genet. Program. Evolvable Mach. 21 (3) (2020) 503–523. .
[22] T. Dutta, S. Bhattacharyya, S. Dey, J. Platos, Border collie optimization, IEEE Access 8 (2020) 109177–109197. .
[23] A.Y. Zebari, S.M. Almufti, C.M. Abdulrahman, Bat algorithm (BA): review, applications and modifications, Int. J. Sci. World 8 (1) (2020) 1.

[24] G. Kicska, A. Kiss, Comparing swarm intelligence algorithms for dimension reduction in machine learning, Big Data and Cognitive Computing 5 (3) (2021) 36.
[25] D. Freitas, L.G. Lopes, F. Morgado-Dias, Particle swarm optimisation: a historical review up to the current developments, Entropy 22 (3) (2020) 362.
[26] N. Rana, M.S.A. Latiff, S.I.M. Abdulhamid, H. Chiroma, Whale optimization algorithm: a systematic review of contemporary applications, modifications and developments, Neural Comput. Appl. 32 (20) (2020) 16245–16277. .
[27] G.Y. Ning, D.Q. Cao, Improved whale optimization algorithm for solving constrained optimization problems, Discrete Dynam Nat. Soc. 2021 (2021). .
[28] V. Kumar, D. Kumar, M. Kaur, D. Singh, S.A. Idris, H. Alshazly, A novel binary seagull optimizer and its application to feature selection problem, IEEE Access 9 (2021) 103481–103496. .
[29] M. Abdel-Basset, L. Abdle-Fatah, A.K. Sangaiah, An improved Lévy based whale optimization algorithm for bandwidth-efficient virtual machine placement in cloud computing environment, Cluster Comput. 22 (4) (2019) 8319–8334. .
[30] R. Salgotra, U. Singh, S. Saha, On some improved versions of whale optimization algorithm, Arabian J. Sci. Eng. 44 (11) (2019) 9653–9691. .
[31] J. Li, H. Lei, A.H. Alavi, G.G. Wang, Elephant herding optimization: variants, hybrids, and applications, Mathematics 8 (9) (2020) 1415.
[32] W. Li, G.G. Wang, A.H. Alavi, Learning-based elephant herding optimization algorithm for solving numerical optimization problems, Knowl. Base Syst. 195 (2020), 105675.
[33] Y. Duan, C. Liu, S. Li, X. Guo, C. Yang, Gradient-based Elephant Herding Optimization for Cluster Analysis, Applied Intelligence, 2022, pp. 1–32. .
[34] A.A. Ismaeel, I.A. Elshaarawy, E.H. Houssein, F.H. Ismail, A.E. Hassanien, Enhanced elephant herding optimization for global optimization, IEEE Access 7 (2019) 34738–34752. .
[35] M.A. Elhosseini, R.A. El Sehiemy, Y.I. Rashwan, X.Z. Gao, On the performance improvement of elephant herding optimization algorithm, Knowl. Base Syst. 166 (2019) 58–70.
[36] J.B. Odili, M.N.M. Kahar, S. Anwar, African buffalo optimization: a swarm-intelligence technique, Procedia Comput. Sci. 76 (2015) 443–448. .
[37] C.P. Igiri, Y. Singh, D. Bhargava, S. Shikaa, Improved African buffalo optimisation algorithm for petroleum product supply chain management, Int. J. Grid Util. Comput. 11 (6) (2020) 769–779. .
[38] A.R. Panhalkar, D.D. Doye, Optimization of Decision Trees Using Modified African Buffalo Algorithm, Journal of King Saud University-Computer and Information Sciences, 2021. .
[39] N.S. Jebaraj, H.R. Keshavan, Hybrid genetic algorithm and african buffalo optimization (HGAABO) based scheduling in ZigBee network, Int. J. Appl. Eng. Res. 13 (5) (2018) 2197–2206. .
[40] P. Singh, N.K. Meena, A. Slowik, S.K. Bishnoi, Modified african buffalo optimization for strategic integration of battery energy storage in distribution networks, IEEE Access 8 (2020) 14289–14301. .
[41] C.P. Igiri, Y. Singh, D. Bhargava, An improved African buffalo optimization algorithm using chaotic map and chaotic-levy flight, Int. J. Eng. Technol. 7 (4) (2018) 4570–4576.
[42] A.W. Mohamed, A.A. Hadi, A.K. Mohamed, Gaining-sharing knowledge based algorithm for solving optimization problems: a novel nature-inspired algorithm, Int. J. Machine Learn. Cybern. 11 (7) (2020) 1501–1529. .
[43] G. Xiong, X. Yuan, A.W. Mohamed, J. Chen, J. Zhang, Improved binary gaining–sharing knowledge-based algorithm with mutation for fault section location in distribution networks, J. Comput. Des. Eng. 9 (2) (2022) 393–405. .
[44] P. Agrawal, K. Alnowibet, A.W. Mohamed, Gaining-sharing knowledge based algorithm for solving stochastic programming problems, Comput. Mater. Continua (CMC) 71 (2) (2022) 2847.
[45] T. Jiang, H. Zhu, G. Deng, Improved African buffalo optimization algorithm for the green flexible job shop scheduling problem considering energy consumption, J. Intell. Fuzzy Syst. 38 (4) (2020) 4573–4589. .
[46] G. Meirelles, B. Brentan, J. Izquierdo, E. Luvizotto, Grand tour algorithm: novel swarm-based optimization for high-dimensional problems, Processes 8 (8) (2020) 980.
[47] D.L. Aditya, D. Fitrianah, Comparative study of fuzzy C-means and K-means algorithm for grouping customer potential in brand LIMBACK, J. Riset Informatika 3 (4) (2021) 327–334. .
[48] Y.A. Wijaya, D.A. Kurniady, E. Setyanto, W.S. Tarihoran, D. Rusmana, R. Rahim, Davies Bouldin Index Algorithm for Optimizing Clustering Case Studies Mapping School Facilities, 2021.
[49] I.F. Ashari, R. Banjarnahor, D.R. Farida, S.P. Aisyah, A.P. Dewi, N. Humaya, Application of data mining with the K-means clustering method and Davies Bouldin index for grouping IMDB movies, J. Appl. Inform. Comput. 6 (1) (2022) 7–15. .
[50] D. Bernard, Clustering Indices, University Paris Ouest Lab Modal'X November, 2017.
[51] S. Al-Janabi, A. Alkaim, E. Al-Janabi, et al., Intelligent forecaster of concentrations (PM2.5, PM10, NO2, CO, O3, SO2) caused air pollution (IFCsAP), Neural Comput. Appl. 33 (2021) 14199–14229, https://doi.org/10.1007/s00521-021-06067-7.
[52] Narjes Nabipour, Amir Mosavi, Eva Hajnal, Laszlo Nadai, Shahaboddin Shamshirband, Kwok-Wing Chau, Modeling climate change impact on wind power resources using adaptive neuro-fuzzy inference system, Eng. Appl. Comput. Fluid Mech. 14 (1) (2020) 491–506, https://doi.org/10.1080/19942060.2020.1722241.
[53] Ghalandari Mohammad, Alireza Ziamolki, Amir Mosavi, Shahaboddin Shamshirband, Kwok-Wing Chau, Saeed Bornassi, Aeromechanical optimization of first row compressor test stand blades using a hybrid machine learning model of genetic algorithm, artificial neural networks and design of experiments, Eng. Appl. Comput. Fluid Mech. 13 (1) (2019) 892–904, https://doi.org/10.1080/19942060.2019.1649196.
[54] K. Guerraiche, L. Dekhici, E. Chatelet, A. Zeblah, Multi-objective electrical power system design optimization using a modified bat algorithm, Energies 14 (13) (2021) 3956.
[55] Shahaboddin Shamshirband, Amir Mosavi, Timon Rabczuk, Narjes Nabipour, Kwok-wing Chau, Prediction of significant wave height; comparison between nested grid numerical model, and machine learning models of artificial neural networks, extreme learning and support vector machines, Eng. Appl. Comput. Fluid Mech. 14 (1) (2020) 805–817, https://doi.org/10.1080/19942060.2020.1773932.
[56] Shahab S. Band, Saeid Janizadeh, Subodh Chandra Pal, Asish Saha, Rabin Chakrabortty, Manouchehr Shokri, Amirhosein Mosavi, Novel ensemble approach of deep learning neural network (DLNN) model and particle swarm optimization (PSO) algorithm for prediction of gully erosion susceptibility, Sensors 20 (19) (2020) 5609, https://doi.org/10.3390/s20195609.
[57] G.G. Wang, S. Deb, X.Z. Gao, L. Dos Santos Coelho, A new metaheuristic optimisation algorithm motivated by elephant herding behaviour, Int. J. Bio-Inspired Comput. 8 (2017) 394–409.
[58] I. Strumberger, M. Minovic, M. Tuba, N. Bacanin, Performance of elephant herding optimization and tree growth algorithm adapted for node localization in wireless sensor networks, Sensors 19 (11) (2019) 2515.