Chapter 18
Evolutionary Machine Learning in Science and Engineering

Jianjun Hu, Yuqi Song, Sadman Sadeed Omee, Lai Wei, Rongzhi Dong, and Siddharth Gianey

Abstract Evolutionary machine learning (EML) has been increasingly applied to solving diverse science and engineering problems due to the global search, optimization, and multi-objective optimization capabilities of evolutionary algorithms and the strong modeling capability of complex functions and processes by machine learning (ML) and especially deep neural network models. They are widely used to solve modeling, prediction, control, and pattern detection problems. Especially EML algorithms are used for solving inverse design problems ranging from neural network architecture search, inverse materials design, control system design, and discovery of differential equations.

18.1 Introduction

Several common fundamental research themes frequently arise in many science and engineering domains such as modeling physical and chemical processes, prediction, pattern classification, abnormality recognition, generation of structures, control, and inverse design. In these problems, the traditional analytical models are increasingly replaced or complemented by data-driven machine ML and especially deep neural network models [70], due to their strong capability to learn complex nonlinear relationships and inherent representations that lead to high-performance prediction models. For example, ab initio crystal structure prediction (CSP) has been a long-time challenging problem due to the expensive first principle calculations needed to evaluate the candidate structures during the search. However, currently, significant progress is being made in learning neural network-based interatomic potentials to speed up the CSP or molecular dynamics simulation process [60]. As a result, ML and especially deep learning (DL) have been transforming almost every discipline of science and engineering.

J. Hu (✉) · Y. Song · S. S. Omee · L. Wei · R. Dong · S. Gianey
Department of Computer Science and Engineering, University of South Carolina, Columbia, SC 29201, USA
e-mail: jianjunh@cse.sc.edu

W. Banzhaf et al. (eds.), *Handbook of Evolutionary Machine Learning*,
Genetic and Evolutionary Computation,
https://doi.org/10.1007/978-981-99-3814-8_18

While deep neural network models are good at modeling, their gradient-descent-based training algorithms can only be applied to differentiable networks. However, there are many science and engineering problems that need strong global (multi-objective) optimization or search capability in vast design space, e.g., of materials and molecules or searching discrete structures such as neural network architectures. In this regard, it is natural to hybridize ML models with evolutionary algorithms to achieve synergistic high performance in problem-solving. Evolutionary computation (EC) has been applied to various stages of ML and DL for model search, feature selection, or hyperparameter tuning. It has played an increasing role in a range of scientific research tasks due to its global search and multi-objective optimization capabilities such as crystal structure prediction or inverse design [5] in which a surrogate performance evaluation model is trained and then used as the objective function in inverse design search. With these complementary roles of ML and evolutionary computation, there emerges the EML paradigm with unique capabilities and applications in diverse disciplines.

EML methods can be understood from different perspectives. From the ML problem point of view, EML has been used in solving supervised learning, unsupervised learning, and reinforcement learning problems. For example, it has been used for clustering [54], classification [106], regression [33, 87], and ensemble learning [40]. From an algorithm and model point of view, EML can be classified into two categories including: (1) EC as ML tools, in which evolutionary algorithms are directly used to solve diverse ML problems such as clustering, classification [41], or regression (2) EC for ML, in which EC is used to improve the model design (such as network architecture as shown in the evolutionary DL [126]), training of ML models, hyperparameter search, feature engineering, and explainability.

Some prominent applications of EML in science and engineering are shown in Fig. 18.1. These applications have some unique characteristics. First, most physical, chemical, and engineering processes are highly nonlinear and are difficult to model explicitly, so neural networks have thus been exploited to model such complex processes. Similarly, the complexity of patterns from these systems has also called for the application of ML and DL in these areas. So both EC as ML and EC for ML have been widely used in science and engineering for clustering, classification, regression, pattern detection, etc. However, there is a special category of application of EML in science and engineering: the (inverse) design problem, which ranges from neural network architecture [15, 66, 111] and parameter design, to partial differential equation discovery [121], inverse materials design [122], neural control system design [84], crystal structure prediction [37], and engineering system design. In these problems, usually, the objective function or performance evaluation is modeled using an ML or neural network model, and then genetic algorithms (GAs) are used to search the candidate solutions in the design space using ML as the objective function. This problem-solving strategy also applies to the problem of discovering physics equations, in which the terms of (partial) differential equations can be discovered by evolutionary algorithms and then assembled either by GAs or by DL.

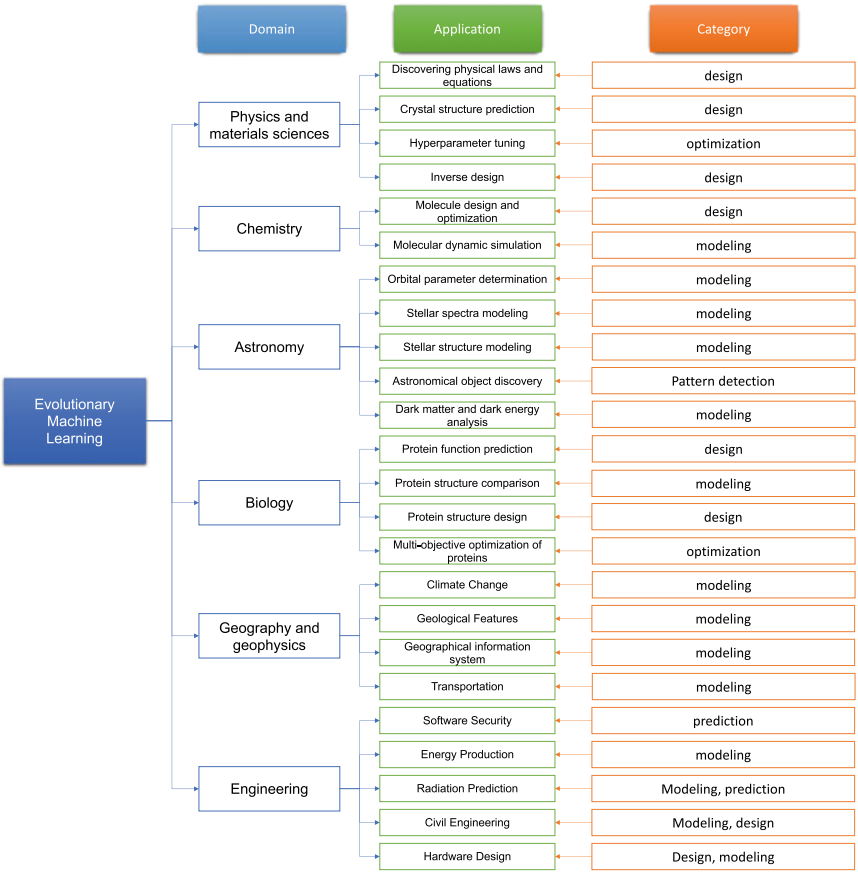

Fig. 18.1 Categories of problems suitable for solving using EML

18.2 Applications of EML in Science and Engineering

EML has gained widespread recognition and has been extensively utilized in diverse fields of science and engineering. In the subsequent sections, we provide a comprehensive overview of the manifold applications of EML in various domains, including physics and materials sciences, chemistry, astronomy, biology, geography, and other branches of engineering. The comprehensive breakdown of the contents pertaining to the applications of EML in these domains is graphically represented in Fig. 18.1.

18.2.1 EML for Physics and Materials Sciences

The application of ML and especially DL in physics and materials science has been accelerating due to the complexity in manually modeling the structure-property relationship, the challenge in ab initio simulation for materials characterization, and the challenge in sampling the huge chemical space. To address these issues, ML and deep neural networks are now commonly used to learn the interatomic potentials for Density Functional Theory (DFT) or molecular dynamics simulation, which can then be combined with global optimization algorithms for crystal structure prediction [8]. Neural networks are also trained with ab initio data to speed up materials property calculation so that they can be used as fast surrogate models for inverse design of materials using evolutionary algorithms [5]. Evolutionary algorithms can also be combined with ML to discover physical laws as represented by partial differential equations [119, 121]. Finally, evolutionary algorithms are also routinely used for training special neural network models due to their gradient-free global optimization capability.

18.2.1.1 Discovering Physical Laws and Equations

Discovering physical laws represented as mathematical models such as partial differential equations (PDE) is one of the most challenging tasks in physics. Data-driven methods have been routinely used for PDE discovery, but they usually require all potential terms to be specified. Xu et al. [119, 121] proposed a novel method, DLGA-PDE for PDE discovery without such constraints. It works by training a DL model of the physical process for generating meta-data and derivatives and then using a GA to search for the combination of such terms. Their method has been shown to achieve good performance in the discovery of the Korteweg-de Vries (KdV) equation, the Burgers equation, the wave equation, and the Chaffee-Infante equation, all with an incomplete term library and even with noisy and limited data. In [26], Chen et al. used an evolutionary strategy to train a discrete feed-forward convolutional neural network model for modeling variational wave functions for correlated many-body quantum systems. They found that networks can converge with high accuracy to the analytically known sign structures of ordered phases. In [120], a robust PDE discovery framework called the robust DL GA (R-DLGA) was proposed, which combines the physics-informed neural network (PINN) and DLGA for potential terms discovery. The terms discovered by DLGA are added to the loss function of the PINN as physical constraints to improve the accuracy of the derivative calculation. The authors [121] further proposed an EML method for DL of parametric partial differential equations from sparse and noisy data. The EML method has also been applied to learn models that balance accuracy with parsimony in classical mechanics and the melting temperature prediction of materials [38]. The EML approach allowed them to discover interpretable physical laws from data based on parsimonious neural networks (PNNs) combined with evolutionary optimization. The EML approach

has also been applied to learn parametric partial differential equations including the Burgers equation, the convection-diffusion equation, the wave equation, and the KdV equation from sparse and noisy data [121], in which the neural network is first trained to calculate derivatives and generate meta-data, which solves the problem of sparse noisy data. Then, the GA is used to discover the form of PDEs and corresponding coefficients.

18.2.1.2 Crystal Structure Prediction

One of the major challenges in materials science is to predict a structure given only its composition. While several crystal structure prediction algorithms based on global optimization and first principle (e.g., Density Functional Theory (DFT)) have been proposed, the computational complexity of such DFT simulations have made it difficult to predict structures for most materials. To address this issue, a series of research combined deep neural networks with GAs for crystal phase determination [8, 37]. In [8], Artrith et al. trained a specialized ML neural network potential using around 1000 first-principles calculations, which can help sample low-energy atomic configurations in the entire amorphous LixSi phase space. The result is comparable to that of ANN trained with extensive molecular dynamics simulations with $\approx 45\,000$ first-principles calculations. Such neural network potentials (including graph neural network potentials) have since been further improved or developed and applied with evolutionary algorithms [17, 28]. In [130], Bayesian optimization is combined with graph neural network (GNN) potential to do crystal structure relaxation. The GNN is then later combined with a GA for de novo crystal structure prediction [31]. Wanzenböck et al. [76] proposed an algorithm that explores the rich phase diagram of TiOx overlayer structures on SrTiO3(110) by combining the covariance matrix adaptation evolution strategy (CMA-ES) and a neural network force field (NNFF) as a surrogate energy model, which dramatically reduces the computational resources needed by DFT simulation. While most neural network potentials are trained before the genetic search of crystal phases, it is also possible to conduct both simultaneously, which is how active learning works. In [91], Podryabinkin et al. proposed a methodology for crystal structure prediction based on the evolutionary algorithm and the active learning of neural network interatomic potentials. Their approach allows for an automated construction of an interatomic interaction model from scratch achieving a speedup of several orders of magnitude. They have benchmarked their algorithms on crystal structure prediction of carbon, high-pressure phases of sodium, and boron allotropes, including those that have more than 100 atoms in the primitive cell, all with satisfactory results. Kang et al. [60] trained a deep neural network interatomic potential model and combined it with a GA for crystal structure prediction. By avoiding the expensive DFT calculations of formation energy and harnessing the speed and accuracy of neural network potentials (NNPs), their algorithm navigates configurational spaces $10^2 - 10^3$ times faster than DFT-based methods. Their SPINNER algorithm has identified more stable phases in many cases than the data-mined template-based method and DFT-based evolutionary algorithm methods.

18.2.1.3 Hyperparameter Tuning

Another major application of EML in physics and materials is hyperparameter tuning of ML models as shown in [109]. Both particle swarm optimization (PSO) and GAs are routinely used to select optimal hyperparameter values autonomously. Tani and Rand evaluated how PSO and GA can improve the performance of their XGBoost ML model for the ATLAS Higgs boson ML challenge (HBC) [3], which represents a typical application of ML algorithms to the field of high energy physics. The task of the HBC is to separate the Standard Model (SM) Higgs boson signal from the large SM background. They showed that compared to using the default hyperparameters, the optimization of the hyperparameter values by GA or PSO improves the sensitivity of the data analysis, by 12-13%, demonstrating that the optimization of hyperparameters is a worthwhile task for data analyses in the field of HEP. In [50], DL techniques are applied to learn the physics of extensive air showers in which the inner structure of the neural network is optimized through the use of GAs.

18.2.1.4 Inverse Design

Another major application of EML in physics is the inverse design in which the ML models are used to learn the relationship between structures and physical property while evolutionary algorithms are used to search the design space. In [32, 46], Comin and Hartschuh combined neural networks with a GA for optimizing spectral-phase shaping of an incident field to achieve second harmonic generation hotspot switching in plasmonic nanoantennas design. They first trained a neural network to predict the relative intensity of the second-harmonic hotspots of the nanoantenna for a given spectral phase and then used a GA to generate a wide range of nanoantenna designs to be fed into the neural network. Taking advantage of the multi-objective optimization feature of the GA, Li et al. [75] applied a multi-objective GA to the optimization of the apertures of the National Synchrotron Light Source II (NSLI-II) Storage Ring. To maintain the diversity of the population of the GA, a K-means clustering algorithm is applied to the population to group individuals into clusters of different fitness levels. Then individuals from the "Good" and "Poor" clusters are mixed with the "Best/elite" individuals to ensure population quality improvement without losing diversity. In [29], Chen et al. proposed an EML approach for physics-guided ML-based inverse design of acoustic metamaterials. They used a multi-layer perceptron (MLP) neural network to map the wave-field-to-wave propagation relationship and then used it as a surrogate model for GA for inverse design of the metabeam. While deep neural networks are widely used as surrogate models, they are not able to tune the output to encompass a range of input states. In this case, a nonlinear symbolic regression model by genetic programming is more desirable, which is used in metamaterials design [5]. In their work, the actual design is implemented using the MLP-based autoencoder network model. The EML-based inverse design has also been applied to designing solid oxide fuel cells (SOFCs) [122], in which a deep neural network is used to map the current density, anode flow rate, cathode flow rate, and temperatures

to outputs that reflect the efficiency and thermal behavior of the SOFC cell. Moreover, the EML has also been used in inverse design of photonics in [52].

There are several other novel applications of EML in physics problems. In [26], Chen et al. used a MLP network to approximate a novel class of variational wave functions for correlated many-body quantum systems. They encode the all-important rugged sign structure of a quantum wave function in a convolutional neural network with discrete output, which is then trained with a gradient-free evolution strategy (ES) algorithm rather than the commonly used back-propagation algorithm. They found that while the stochastic gradient descent (SGD) algorithm is better for optimizing continuous functions, the ES method is the better choice for optimizing the variational wave function while SGD is no longer applicable. In [116], a multi-objective GA is combined with deep neural network models for improving the performance and durability of direct internal reforming solid oxide fuel cells.

18.2.2 EML for Chemistry

Evolutionary algorithms (EAs) are generic, population-based, metaheuristic optimization methods. The mechanisms by which EAs operate are inspired by biological evolutionary operations such as selection, mutation, recombination, and reproduction. Combined with ML, GAs provide a novel tool for the investigation of molecule design, optimization, and molecular dynamic simulation.

18.2.2.1 EML for Molecule Design and Optimization

The rise of machine intelligence provides a grand opportunity to expeditiously design and discover novel molecules through smart search. The discovery of new functional molecules has led to many technological advances and remains one of the most critical approaches to overcoming technical problems in various industries, such as those in organic semiconductors, displays, and batteries.

Evolutionary design has gained significant attention as a useful tool to accelerate the design process by automatically modifying molecular structures to obtain molecules with the target properties. However, devising a way to rapidly evolve the molecule while maintaining its chemical validity is a challenge. Kwon et al. [68] proposed a GA along with RNN and DNN models that were used to evolve the fingerprint vectors of seed molecules. The RNN decoder reconstructed chemically valid molecular structures in the SMILES format from the evolved fingerprint vectors without resorting to predefined chemical rules. The method employed DL models to extract the inherent knowledge from a database of materials and is used to effectively guide the evolutionary design. Designing a new therapeutic drug can be a time-consuming and expensive process. It also enables rapid discovery of new drug candidates by performing intelligent searches in a wide molecular structure space. In [2], Abouchekeir et al. proposed a new approach called adversarial deep

evolutionary learning (ADEL) to search for novel molecules in the latent space of an adversarial generative model and keep improving the latent representation space. In [125], Yoshikawa et al. proposed a new population-based approach using a grammatical evolution named ChemGE that can update a large population of molecules concurrently and ChemGE succeeded in finding hundreds of candidate molecules whose affinity for thymidine kinase is better than that of known binding molecules in a database (DUD-E). Li et al. [74] proposed a deep evolutionary learning (DEL) process that integrates a fragment-based deep generative model and multi-objective evolutionary computation for molecular design, which can generate promising novel molecular structures.

Optimizing molecules for desired properties is a fundamental yet challenging task in chemistry, material science, and drug discovery. In [27], Chen et al. developed a novel algorithm for optimizing molecular properties via an Expectation Maximization (EM) like an explainable evolutionary process. They showed that the evolution-by-explanation algorithm is 79% better than the best baseline in terms of a generic metric combining aspects such as success rate, novelty, and diversity.

Evolutionary algorithms have found increasing applications in both the discovery and optimization of novel molecular structures. Artificial evolutionary methods, such as GAs, can not only explore large and complex search spaces very efficiently, but also can be applied to the identification and optimization of new molecules faster than pure physical experiments. ML models can enhance the suitability of experimentally measured molecules to accelerate the discovery of useful and novel molecules in a broad composition or property space. For example, Tu C. Le and Nhiem Tran [69] reviewed how GAs have been used to solve optimization problems in computational drug design including catalyst discovery and optimization. They also describe the use of both experimental and computational fitness functions to evolve materials into promising areas of catalyst space. Among these applications, neural networks have been used widely as in silico fitness functions based on these neural structure-property models.

18.2.2.2 EML for Molecular Dynamic Simulation

Molecular dynamics (MD) has become a powerful tool for studying biophysical systems due to increased computing power and the availability of software. Efficient computational strategies for the targeted generation and screening of molecules with desired therapeutic properties are therefore urgently required.

Because of the negative environmental impact of damaging organic solvents and the high cost of chemical waste disposal, the search for alternative, renewable solvents is a top priority in the chemical industry. Zhong et al. [128] facilitated the development of optimized potentials for liquid simulation (OPLS)-based force field (FF) parameters for eight unique deep eutectic solvents (DESs) based on three ammonium-based salts and five HBDs at multiple salt-HBD ratios. DESs are a class of solvents often composed of ammonium-based chloride salts and a neutral hydrogen bond donor (HBD) at specific ratios. These cost-effective and environmentally friendly

solvents have seen significant growth in multiple fields, including organic synthesis, and in materials and extractions because of their desirable properties. Coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus type 2 (SARS-CoV-2) has led to a global pandemic. DL techniques and molecular dynamics (MD) simulations are two mainstream computational methods for studying the geometric, chemical, and structural features of proteins and guiding protein structures. Sun et al. [107] introduced the latest progress of the DL-based molecular dynamic simulation approaches in structure-based drug design (SBDD) for SARS-CoV-2 which could address the problems of protein structure and binding prediction, drug virtual screening, molecular docking, and complex evolution. MD can also help researchers develop new molecular materials in biomaterials science. Collagen is the most abundant structural protein in humans, providing crucial mechanical properties, including high strength and toughness, in tissues. In [61], Khare et al. developed a general model using a GA within a DL framework to design collagen sequences with specific Tm values. They discovered that the number of hydrogen bonds within collagen calculated with molecular dynamics (MD) is directly correlated to the experimental measurement of triple-helical quality.

18.2.3 EML for Astronomy

The fundamental advantage of EML is that they deliver high-quality results even when computational resources are constrained. The so-called comprehensive optimization methods are primarily constrained by the scale of the problem or data since the field of astronomy often involves optimizing problems of great complexity or processing enormous amounts of data. For this reason, EML has been used in various significant applications in astronomy such as orbital parameter determination, stellar spectra modeling, stellar structure modeling, planet search, gamma emission analysis.

18.2.3.1 Orbital Parameter Determination

Due to the sheer number of unknown parameters involved in the problem of finding orbital parameters, it is considered a highly difficult task. As a result, in the absence of an efficient approach, one is forced to either accept an extremely coarse-grained parameter space scan or limit the search space by setting the ranges of certain parameters. Wahde [112] proposed a method based on GA for efficiently searching vast space of possible orbits. The goal of this paper was to evaluate the effectiveness of using a GA-based method to determine the orbital parameters of interacting galaxies provided photometric observations and systemic velocities of the pair of galaxies. Later Wahde and Donner [113] extended the simulation part of this GA-based approach to investigate the impact of past interactions between the NGC5195 and the Messier M51 galaxies. In another work, Theis and Kohle [110] showed that

if adequate data are provided, their GA-based method can reliably calculate orbital parameters. The use of a GA theoretically enables a uniqueness test of a preferred parameter combination and so the authors use it on the parameter region determined after the fast restricted N-body method. Cantó et al. [21] designed a modification of the canonical GA to observe the orbital parameters of the planets orbiting 55 Cancri. The GA is predominantly used here to maximize a function where traditional methods are ineffective.

18.2.3.2 Stellar Spectra Modeling

Modeling a good fit of cosmic stellar spectra is a very challenging objective. However, this task is very crucial as a wide range of stellar attributes can be inferred from this. The first known evolutionary computing-based analysis of stellar spectra was performed by Metcalfe [79]. The work proposed employing a GA to compare the observed light curves to those produced by theoretical models to determine the properties of binary stellar systems. In this work, GA was used to randomly populate the defined parameter space, which also allowed the trial parameter sets to evolve over time. The optimal set of parameters and the mean set of parameters have very negligible differences after 100 generations. Mokiem et al. [82] developed a parallelized GA which served as the foundation for an autonomous fitter of the spectra of massive stars with stellar winds. With the utilization of a rapid performance stellar atmosphere code named FASTWIND [92] and a fitting method based on GA named PIKAIA [25], a fast and efficient method for automating the fitting of the continuum normalized spectra of O^- and early B^- type stars with stellar winds is described in this paper. The GA-based routine PIKAIA is employed for parameter optimization of FASTWIND. PIKAIA optimizes a population till a predetermined number of generations is reached rather than until a specified criterion is met. PIKAIA is also used on other several noteworthy papers. One such important application can be found in the work by Baier et al. [14]. They combined the radiative transfer code DUSTY [58] with PIKAIA to notably enhance the spectral fit of the dust spectra of AGB stars. PIKAIA, which is based on the evolutionary natural selection process, seeks to maximize the function $g(\lambda) = [F(\lambda) - F_m(\lambda)]^{-2}$, where $F(\lambda)$ is the observed spectrum and $F_m(\lambda)$ is a model spectrum computed using DUSTY.

18.2.3.3 Stellar structure Modeling

Another important application of EML in astronomy is stellar structure modeling. The internal structures of stellar objects vary depending on their classifications and ages, reflecting the components they are made of and how they transfer energy. Metcalfe and Charbonneau [80] obtained a number of intriguing physical discoveries for stellar structure modeling of white dwarf stars, thanks to the effective, concurrent exploration of parameter space enabled by GA-based numerical optimization. The authors also use the GA-based routine PIKAIA [25] (described in the

paragraph above) and re-implement it as a fully parallel routine to provide an objective estimate of the globally optimal parameters for a particular model versus an observational dataset. A total of five parameters were allowed to evolve by the parallel GA for obtaining structural and physical details about the white dwarf stars. Zhang et al. [127] calculated stellar effective temperatures and identified angular parameters using a stochastic PSO on known stellar flux data in specific bands. The system's input settings were first set. Second, each particle's fitness value was determined. The fitness values of each particle with the prior best predictions were then contrasted. Following the generation of the new particle, other particles' positions and velocities were updated. When a specific stopping criterion is met, the stochastic PSO terminates. Another work describes a novel approach based on GA for estimating the age and relative contribution of various stellar populations in galaxies [9]. Using charge-coupled device (CCD) images in the U, B, V, R, and I bands, the authors apply this technique to the barred spiral galaxy NGC 3384. Using the hypothesis that just two stellar populations, each with a different color, age, metallicity, etc., are responsible for the observed light from a galaxy, the GA is used to solve the equation set that describes the relationships characterizing the two stellar populations' mixing-tracks found from [1].

18.2.3.4 Astronomical Object Discovery

Even though there are a great number of astronomical objects, we have only discovered and examined a small fraction of them. This necessitates the creation of algorithms for efficiently finding astronomical objects. Nesseris and Shafieloo [86] developed a null test for the cosmological constant model using the so-called Om statistic in combination with GAs to recreate the expansion history of the universe in a model-independent way. In this paper, the GA is applied on the SNIa dataset (with the selected execution parameters) [99] to find a solution for the "distance modulus" term of its fitness function. Based on evolutionary optimization of the classifiers, Wierzbiński et al. [117] created an effective and precise classifier for cosmic objects that are mostly used to discover the best parameters for the voting classifiers to categorize stellar spectra of stars, quasars, and galaxies. With their default parameters as a starting point, the authors trained a collection of 21 classifiers. Then, they optimized the hyperparameters using GA. Cassisi [23] used an evolutionary algorithm in which the equations for convective mixing and nuclear burning are solved using a single common scheme to undertake the first complete evolutionary computations of stars undergoing "He flash mixing". For the evolutionary calculations, the authors employ an evolutionary code from [103]. Joseph et al. [59] used GA to search through four nonlinear parameters of each planet to fit a full Keplerian orbit. Chan et al. [24] used a GA for locating stars formed after supernovae explosion and eventually backtracking to stars formed after the big bang. As applying a complete search is computationally very expensive, the authors employed a GA with a local search boosted with random initial solutions. In a study by Geyter et al. [34], the SKIRT [12], a Monte Carlo radiative transfer code designed to examine the impact of dust absorption and

scattering on the gas and star kinematics of dusty galaxies, its output is optimized using the GA library GAlib [114] by searching through the vast model parameter space.

18.2.3.5 Dark Matter and Dark Energy Analysis

Dark energy makes up around 68% of the universe and dark matter about 27%. They are predominantly responsible for the bulk of galaxies and galaxy clusters as well as the large-scale organization of galaxies. So they are very significant to understand the formation of the universe. Bogdanos and Nesseris [18] employed GAs to study standard SNIa data [67] to extract model-independent restrictions on the evolution of the dark energy equation of state. With the selected execution settings, the GA is applied to the original SNIa dataset to produce a solution for the reduced distance modulus. Ruiz et al. [95] applied PSO to analyze merger trees obtained from a common Lambda Cold Dark Matter N-body simulation and the Semi-analytic Model of Galaxy Formation (SAM). The PSO is mainly used here to calculate the best possible set of SAM parameters. Moster et al. [83] devised a reinforcement learning approach to compute the galaxy properties for dark matter haloes and train the parameters using a PSO technique. The authors compute the galaxy properties for all haloes for a specific set of weights and biases and then produce mock statistics for GalaxyNet, which is trained using a reinforcement learning approach. These statistics are compared to the observations to determine the model loss, which is minimized using PSO.

18.2.4 EML for Biology

The protein structure design problem is one of the most exciting challenges of modern computational biology. Because of its scientific complexity, research on understanding the function of proteins, and studying the relationship between amino acid sequences and protein structures is very difficult. ML-guided evolution is a new paradigm for a biological design that enables optimizing complex procedures. And the multi-objective evolutionary algorithm is introduced as it can deal with several functions when designing protein structures.

18.2.4.1 Protein Function Prediction

Protein function prediction methods are techniques that bioinformatics researchers use to understand the biological or biochemical roles of proteins. Proteins often function poorly when used outside their natural contexts and directed evolution can be used to engineer them to be more efficient in new roles. Wu et al. [118] incorporate ML into the protein's directed evolution workflow, to reduce the experimental effort and to

explore the sequence space encoded by mutating multiple positions simultaneously. They have validated their approach on a large published empirical fitness landscape and demonstrated that ML-guided directed evolution finds variants with higher fitness than other evolution approaches.

18.2.4.2 Protein Structure Comparison

Identifying structural similarities is essential for assessing the relationship between structure and function in proteins. Szustakowski and Weng [108] developed a structure alignment algorithm using GA for three-dimensional structures of proteins. They first align the proteins' cores, as represented by their secondary structure elements, by minimizing the difference of distance matrices using a GA. And then extend the alignment to include any positions in loops or turns deemed equivalent in a convergent process. Carr et al. [22] developed a new approach to structural comparison by using a Multimeme evolutionary algorithm. In a Multimeme algorithm, an individual is composed of its genetic material (that represents the solution to the problem being solved) and its memetic material (that defines the kind of local searcher to use). During crossover, both genetic and memetic transmission will be done. In [11], Bacardit et al. used GA to design automated procedures to reduce the dimension of protein structure prediction datasets by simplifying how the primary sequence of a protein is represented. Reducing the size of the alphabet used for prediction from twenty to just three letters resulted in more compact and human-readable classifiers. And the loss of accuracy accrued by this substantial alphabet reduction is not statistically significant compared to the full alphabet.

18.2.4.3 Protein Structure Design

ML-guided directed protein structure design enables optimization of complex functions. Pegg et al. [90] developed a GA for structure-based de novo design. They use molecular interactions evaluated with docking calculations as a fitness function to reduce the search space. Durrant et al. [39] developed a protein inhibitor design algorithm that uses a growth strategy to build the core scaffold, molecular fragments are added at random to this scaffold. An evolutionary algorithm is then used to evaluate the scores of each population member, and the best ones become founders of the subsequent generation.

18.2.4.4 Multi-objective Optimization of Proteins

Using only one energy function is insufficient to characterize proteins because of their complexity. Multi-objective algorithms can provide a better protein with less computational resource requirement than DL methods. Brasil et al. [19] developed a multi-objective evolutionary algorithm for ab initio protein structure prediction

without using any earlier knowledge from similar protein structures. Gao et al. [44] adopted a solvent-accessible surface area into a multi-objective evolutionary algorithm with three objective functions to improve protein structure prediction accuracy and efficiency. Traditional multi-objective algorithms cannot obtain the desired solution because the selection pressure decreases as the number of objectives increases. To address this problem, Lei et al. [71] proposed a many-objective evolutionary algorithm with four types of objectives to alleviate the impact of imprecise energy functions for predicting protein structures.

18.2.5 EML for Geography and Geophysics

Process in evolutionary ML has contributed to significant advances in geography and geophysics research, such as climate change [94], geological features, geographical information systems, and transportation.

18.2.5.1 Climate Change

Recently, one of the biggest problems confronting humanity is climate change. Storms, droughts, fires, and flooding have become stronger and more frequent [94]. The main strategies for addressing climate change include reducing greenhouse gas (GHG) emissions and preparing for resilience and disaster management, named mitigation and adaption respectively. Mitigation of GHG emissions requires changes to electrical systems, transportation, buildings, industry, and land-use. Many experts are exploring how to use ML methods to tackle the issue of climate change. Considering both adaption and mitigation response to climate change, Paton et al. [89] incorporated GHG emissions into the multi-objective evolutionary algorithm (MOEA). The application of this method in Adelaide, Australia's southern water supply system has illustrated the framework's useful management implications. For spatially allocating land-use, analyzing climate change impacts may be a useful and fundamental long-term adaptation strategy. Joo Yoon and co-workers[124] utilized multi-objective GA to identify climate adaptation scenarios based on existing extents of three land-use classes in South Korea. Specifically, five objectives were established for predicting climate change impacts and regional economic conditions: (1) minimization of disaster damage, (2) existing land-use conversion, (3) maximization of rice yield, (4) protection of high species richness areas, and (5) economic value. This method showed better performance than other spatial land-use compositions for all adaptation objectives. Climate change trends have already affected many ecosystems, such as species range and diversity, and this effect is different and varies from place to place. Rezayan et al. [97] provided an optimal combination of the common species distribution models (SDMs), and employed a GA to model the climate change effects on the spatial distribution of *Quercus brantii* in the west of Iran.

18.2.5.2 Geological Features

Several attempts have been made to predict and analyze geographic properties by EML methods, such as earthquake hypocenter location [64, 96, 102], surface water reservoir control [100], ocean wave height [115], and magnetic anomalies [16].

In the geophysical research area, the determination of reliable and accurate earthquake hypocenters is a crucial task. In 1993, Sambridge and Gallagher [102] published a paper in which they used GAs to predict earthquake hypocenter location, which can refine a population of hypocenters collectively by exploiting information from the group as a whole, rather than relying on only local information about a single hypocenter. Kim et al. [64] applied a GA and a two-point ray tracing method [63] to solve non-uniqueness problems in determining reliable hypocentral parameters including latitude, longitude, source depth, and origin time. With the increasing risk of flood and drought impacts and the changing water allocation requirements among complex users, an efficient multipurpose reservoir management strategy is critical. Salazar et al. [100] carried out a diagnostic assessment framework of the surface water reservoir control for the Conowingo reservoir in the Lower Susquehanna River Basin, Pennsylvania, USA. Specifically, they use Evolutionary Multi-Objective Direct Policy Search (EMODPS) [48] as the decision analytic framework where reservoirs' candidate operating policies are represented using parameterized global approximators. And then, they use multi-objective evolutionary algorithms to optimize those parameters for discovering the Pareto approximate operating policies. To predict the heights of ocean waves accurately and quickly, which is an important problem in marine detection and warning, Wang et al. [115] reported a hybrid Mind Evolutionary Algorithm-BP neural network strategy (MEA-BP). It can avoid early convergence and improve the prediction accuracy by combining the local searching capabilities of the BP neural network and the global searching ability of the MEA. Their experiments cover a wide range of geographical locations (from 12 observation points across two geographically distinct regions - the Bohai Sea, and the Yellow Sea) and different weather (from Jan 1, 2016, to Dec 31, 2016) and show faster running time and high prediction accuracy. Balkaya and co-workers [16] demonstrated the differential evolution algorithm for 3D nonlinear inversion of total field magnetic anomalies caused by vertical-sided prismatic bodies. Li et al. [72] used an artificial neural network (ANN) to capture ionospheric spatiotemporal characteristics with a powerful capacity and capability and used a GA to improve the ANN's learning efficiency for predicting ionospheric peak parameters including foF2 and hmF2.

18.2.5.3 Geographical Information Systems

GAs have recently generated much interest in the field of geographical information system (GIS) including optimal location search [73], and land partitioning [35]. Optimal location search is usually required in many urban applications for establishing one or more facilities. When it involves multiple sites, various constraints, and multiple objectives, the search task is very complex. In work [73], it demonstrated that

GAs can be used with GIS to effectively solve the spatial decision problems for n facilities. Land partitioning is a basic process of land consolidation [105], it involves the subdivision of land into smaller sub-spaces subject to several constraints [36]. Demetriou et al. [35] proposed the GIS and GAs integrated model: Land Parceling System (LandParcelS) which automates the land partitioning process by designing and optimizing land parcels according to their shape, size, and value.

18.2.5.4 Transportation

Bicycle-sharing systems have become an important part of urban transportation systems [43]. To increase the number of bicycles available for rent and improve profits, it is better to collect bicycles in the evening and redistribute them to the main stations. This relies on the model that accurately forecasts rental demand. In [45], Gao et al. presented a moment-based rental prediction model by a fuzzy C-means (FCM)-based GA with a back propagation network (BPN) with more than ten factors, including the date, time, weather (e.g., temperature, humidity, and wind speed) and season, all of which make different contributions to the final demand. Firstly, they use the unsupervised FCM-based GA method to pre-classify historical rental records into groups. Next, the classification results are fed into a BPN predictor which is trained using these categorized records. After training, the BPN predictor can predict the demand at future moments.

18.2.6 EML for Engineering

Evolutionary algorithms are being used to enhance and optimize traditional ML systems. The quick convergence and flexibility that comes with evolutionary algorithms make their performance so useful in so many different areas. The use of evolutionary algorithms with ML is growing, and it is being used in many engineering areas such as software security, energy production, radiation prediction, civil engineering, and hardware design.

18.2.6.1 Software Security

Cybersecurity and preventing unwanted attacks on computer networks is a very important and constantly changing field. Intrusion detection systems (IDS) are being used to keep malicious software out and are constantly being developed. With the addition of ML algorithms, these IDSs can classify network traffic as normal or not. Hosseinei and Zade [53] made a novel hybrid intrusion detection method that has two phases, a feature selection phase, and an attack detection phase. In the feature selection phase, a GA is modified to have a new multi-parent crossover operator and combined with a support vector machine (SVM) to improve performance and help

reduce the dimensions and select relevant features. For the attack detection phase, an ANN was used to detect any attacks. To train the ANN, a hybrid system was created using hybrid gravitation search and PSO algorithms. Al-Yaseen et al. [4] proposed another IDS that involved the use of a wrapper-based differential evolution algorithm for feature selection and an extreme learning machine as the classifier. The differential evolution algorithm is similar to a GA, but mutation occurs first before crossover. The feature selection from this algorithm is fed into an ML algorithm to classify and detect intruders. In [51], Halim et al. proposed a similar IDS system that used a GA for feature selection as well as an ML algorithm-based prediction system. Another aspect of cyber security that is often developing is the detection of spam and phishing websites. Spammers are creating new techniques that dodge filters and phishing attacks are becoming a larger problem for web security. In [42], Farris et al. proposed a new detection system for spam using a GA in combination with a random weight network. This new system can better handle the massive data flow that needs to be processed to be able to accurately filter out spam. The system uses a GA to find possible feature subsets and an RNW as the base classifier. Ali and Ahmed [6] created a new hybrid phishing website prediction model. The model uses a GA for feature selection, figuring out the most important features and providing proportional weight to these features for optimal results, and applies it to a deep neural network to accurately predict which websites are phishing sites. Detecting where future attacks on software could come from is also very important to the security of that software. These are done using a vulnerability prediction model (VPM). Sahin et al. [98] designed a new model to predict vulnerabilities in software, applying symbiotic GA and DL methods. The proposed VPM used a deep neural network-based symbiotic GA to predict where the software may be lacking in security. The method used two versions of GA and a population-based dominance mechanism to identify the dominant-feature representations.

18.2.6.2 Energy Production

Energy production and consumption are both very necessary for this world. The shift to more natural energy sources has already started, but to see if they can even keep up with what a place or country needs, the amount of renewable energy produced needs to be optimized and measured, and the energy demand needs to be predicted. Prediction systems are developed to help with this, but the feature selection process can be improved with the help of evolutionary approaches. Zhou et al. [129] created a prediction system of photovoltaic cells to see solar energy production. The system utilized a GA and a customized similar day analysis based on an extreme learning machine to predict the photovoltaic output. In [101], Salcedo-Sanz et al. focused on renewable energy uses and developed a new system that used a coral-reef optimization algorithm with a substrate layer for feature selection. Then an extreme learning machine was used to predict renewable energy production. Hu and Chen [55] were looking more specifically at wind energy and wind speed prediction. The proposed model included a differential evolutionary algorithm in conjunction with a Long

Short Term Memory (LSTM) to accurately predict wind speeds and wind energy production. Neshat et al. [85] also used an LSTM in combination with six different evolutionary algorithms to test them all for wind speed and wind energy production predictions. Prediction systems for the demand of places are also being developed and enhanced using evolutionary algorithms. Mason et al. [78] created a new system that applied a Covariance matrix adaptation evolutionary strategy to train neural networks to predict the energy demand for Ireland. It resulted in accurate predictions with fast convergence times. Seyedzadeh et al. [104] similarly made an energy consumption prediction model for non-domestic buildings. A gradient-boosted regression tree model was used for prediction, and an evolutionary algorithm was used to optimize the model by adjusting the hyperparameters.

18.2.6.3 Radiation Prediction

Ghimire et al. [47] proposed a self-adaptive differential evolutionary extreme learning machine (ELM) to predict daily solar radiation for solar-rich cities. For the feature selection for the predictors of solar radiation, a swarm-based ant colony algorithm was used. The learning algorithm used was optimized from a standard ELM: a self-adaptive differential evolutionary extreme learning machine. This system was compared to others used to accomplish the same task, and it performed on par, if not more accurately for predicting solar radiation forecasts. In [62], Kilic et al. were looking at global solar radiation prediction. They implemented a hybrid ANN system, using an evolutionary algorithm for feature selection and to help train the ANN. The hybrid system was compared to usual ML algorithms for the prediction of radiation, including an ANN, SVM, and a DL model. Guijo-Rubio et al. [49] did something similar, using an evolutionary algorithm to train and evolve an ANN to predict solar radiation. Marzouq et al. [77] utilized a GA for the selection of critical inputs for the ANN to predict solar radiation. This new evolutionary ANN model outperformed usual algorithmic models for solar radiation prediction. Amiri et al. [7] took the solar radiation prediction further, predicting on a tilted plane. The evolutionary algorithm was used to find the proper topology, and the study was conducted in two ways, assuming the satellites where the data was pulled from had the same tilt angle, and assuming the angles could be different.

18.2.6.4 Civil Engineering

Inducing blasts is a common occurrence in rock mining and construction. Being able to predict the vibrations of these blasts and the impact they will have on the surroundings is key to the safety of the operation and the workers themselves. When researching the vibrations, the goal is to measure the peak particle velocity as a base parameter. Chen et al. [30] proposed new solutions to the problem, combining various evolutionary algorithms with either an SVM or an ANN. A firefly algorithm, GA, PSO algorithm, and a modified firefly algorithm were all hybridized with both

an SVM and an ANN in two separate systems, and the results of all were compared. The systems were all accurate at predicting peak particle velocity (PPV), with the modified firefly algorithm support vector regression (SVR) system performing the best. Azimi et al. [10] did something similar, focusing on using a GA to optimize an ANN to predict blast vibrations for quarry mining. When constructing different reinforced concrete structures, beam-column joints are an important point in the construction, often being the first place to shear and break off. Yaseen et al. [123] predicted the shearing of these joints with the use of a GA combined with a deep neural network. The GA was used for the input selection during the modeling phase, and the DLNN was used for the prediction phase: to identify any structural problems. Huang et al. [56] also predicted the shearing of steel fiber-reinforced concrete columns. Two hybrid systems were created, combining a GA with an ANN and a PSO algorithm with an ANN. In [20], Cai et al. used a GA hybridized with a back-propagation neural network to predict the capacity reinforced concrete beams had to flex in the structure. Moayedi et al. [81] created a prediction system for ultimate bearing capacity, to see the possible load a bearing can handle. Many systems were developed, including a hybrid GA and ANN system, and a PSO and ANN system to predict the bearing capacity.

18.2.6.5 Hardware Design

The optimal design for hardware and technologies is a big part of engineering and production. Owoyele et al. [88] proposed an automated ML GA as the system for optimizing the design of engines. Using computational fluid dynamic simulations, the system can optimize engines with a hands-off approach. In [13], Bagheri et al. designed a system that would optimize the design of piezoelectric energy harvesters using simulations. The system included a NN which was trained by a GA to optimize the prediction of the optimal design. Technology is becoming more and more advanced, and in so-called advanced technologies, standard cell layouts are done manually for the design of their nodes. Ren and Fojtik [93] proposed a system using reinforcement learning in combination with a GA to help predict the optimal layout of the standard cells. Materials that are for the absorption of microwaves are important to technologies related to stealth. Huang et al. [57] proposed a large mutation GA to optimize the honeycomb meta-structure to have absorption coverage over broadband microwaves as well as have successful mechanical resistance. The deployment of new technology is also very important, trying to efficiently upgrade a system with minimal downtime. Ko [65] designed a system using reinforcement learning hybridized with a GA to create a system to deploy a new wireless charging tram system optimally.

18.3 Discussion

EML has experienced a significant upward trend and facilitates fundamental research in various science and engineering fields, including physics, chemistry, astronomy, biology, geography, and engineering. The technology has proven invaluable in the development of machine learning algorithms that can evolve and adapt in response to changing data and environmental conditions, enabling researchers to gain a deeper understanding of complex systems. To illustrate the scope and impact of EML-related research, we have included a summary of related works in Fig. 18.1. We hope this provides a glimpse into the exciting possibilities that EML can offer to researchers across various disciplines.

Despite the promising advancements in EML, there are still several challenges that need to be addressed. One of the key challenges is the optimization of EML algorithms, which requires the use of efficient search techniques to identify the best solutions in large and complex search spaces. Another challenge is the interpretability of EML models, which can be difficult due to the complex and nonlinear relationships between variables. Additionally, EML requires significant computational resources, including high-performance computing and large-scale data storage, which can be costly and difficult to maintain. Furthermore, EML models may suffer from overfitting or underfitting, leading to poor generalization and limited applicability in real-world scenarios.

Challenges also represent opportunities, EML can also be applied to a wide range of domains, including finance, business, healthcare, and others. In these domains, EML can be used to develop predictive models that can inform decision-making and improve efficiency. Additionally, EML can be used for optimization tasks, such as scheduling and resource allocation, to improve performance and reduce costs. Furthermore, EML can be used in combination with other advanced deep learning techniques, such as transformer and diffusion models, to improve model performance and achieve more accurate predictions. By leveraging the strengths of both approaches, researchers can develop more robust and effective algorithms and models.

18.4 Conclusion

The combination of evolutionary computation and ML has unique advantages in solving challenging modeling and design problems arising in diverse fields of sciences and engineering. This is due to the powerful gradient-free evolutionary search and the strong modeling capability of ML and especially deep neural network models. In this chapter, we have shown that the EML paradigm has been widely applied in physics, materials science, chemistry, astronomy, life sciences, geology, and engineering, ranging from black-box modeling, analytical modeling, data-driven

differential equation learning, inverse design, hyperparameter optimization, training of deep neural network models, multi-objective optimization and design using ML surrogate models.

18.5 Contribution

Conceptualization, J.H.; resources, J.H.; writing–original draft preparation, J.H., Y.S. L.W., S.O., R.D., S.G.; writing–review and editing, J.H; visualization, J.H. and Y.S.; supervision, J.H.; funding acquisition, J.H.

Acknowledgements The research reported in this work was supported in part by National Science Foundation under the grant 2110033, 1940099, and 1905775. The views, perspectives, and content do not necessarily represent the official views of the NSF. We appreciate the help of Dylan Johnson for proofreading the manuscript.

References

1. Abdel-Hamid, H., Notni, P.: Stellar population analysis from broad-band colours. Astron. Nach. **321**(5–6), 307–314 (2000)
2. Abouchekeir, S., Tchagang, A., Li, Y.: Adversarial deep evolutionary learning for drug design. In: 2021 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), pp 01–09. IEEE (2021)
3. Adam-Bourdarios, C., Cowan, G., Germain, C., Guyon, I., Kégl, B., Rousseau, D.: The higgs boson machine learning challenge. In: NIPS 2014 Workshop on High-Energy Physics and Machine Learning, pp. 19–55. PMLR (2015)
4. Al-Yaseen, W.L., Idrees, A.K., Almasoudy, F.H.: Wrapper feature selection method based differential evolution and extreme learning machine for intrusion detection system. Pattern Recogn. **132**, 108912 (2022)
5. Alderete, N.A., Pathak, N., Espinosa, H.D.: Machine learning assisted design of shape-programmable 3d kirigami metamaterials. NPJ Comput. Mater. **8**(1), 1–12 (2022)
6. Ali, W., Ahmed, A.A.: Hybrid intelligent phishing website prediction using deep neural networks with genetic algorithm-based feature selection and weighting. IET Inf. Secur. **13**(6), 659–669 (2019)
7. Amiri, B., Gómez-Orellana, A.M., Gutiérrez, P.A., Dizène, R., Hervás-Martínez, C., Dahmani, K.: A novel approach for global solar irradiation forecasting on tilted plane using hybrid evolutionary neural networks. J. Clean. Prod. **287**, 125577 (2021)
8. Artrith, N., Urban, A., Ceder, G.: Constructing first-principles phase diagrams of amorphous li x si using machine-learning-assisted sampling with an evolutionary algorithm. J. Chem. Phys. **148**(24), 241711 (2018)
9. Attia, A.-F., Ismail, H.A., Selim, I.M., Osman, A.M., Isaa, I.A., Marie, M.A., Shaker, A.A.: Stellar population analysis of galaxies based on genetic algorithms. Chinese J. Astron. Astrophys. **5**(4), 347 (2005)
10. Azimi, Y., Khoshrou, S.H., Osanloo, M.: Prediction of blast induced ground vibration (bigv) of quarry mining using hybrid genetic algorithm optimized artificial neural network. Measurement **147**, 106874 (2019)

11. Bacardit, J., Stout, M., Hirst, J.D., Sastry, K., Llora, X., Krasnogor, N.: Automated alphabet reduction method with evolutionary algorithms for protein structure prediction. In: Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation, pp. 346–353 (2007)
12. Baes, M., Verstappen, J., De Looze, I., Fritz, J., Saftly, W., Pérez, E.V., Stalevski, M., Valcke, S.: Efficient three-dimensional nlte dust radiative transfer with skirt. Astrophys J. Suppl. Ser. **196**(2), 22 (2011)
13. Bagheri, S., Nan, W., Filizadeh, S.: Application of artificial intelligence and evolutionary algorithms in simulation-based optimal design of a piezoelectric energy harvester. Smart Mater. Struct. **29**(10), 105004 (2020)
14. Baier, A., Kerschbaum, F., Lebzelter, T.: Fitting of dust spectra with genetic algorithms-i. perspectives and limitations. Astron. Astrophys. **516**, A45 (2010)
15. Baldominos, A., Saez, Y., Isasi, P.: On the automated, evolutionary design of neural networks: past, present, and future. Neural Comput. Appl. **32**(2), 519–545 (2020)
16. Balkaya, Ç., Ekinci, Y.L., Göktürkler, G., Turan, S.: 3d non-linear inversion of magnetic anomalies caused by prismatic bodies using differential evolution algorithm. J Appl. Geophys. **136**, 372–386 (2017)
17. Behler, J.: Four generations of high-dimensional neural network potentials. Chem. Rev. **121**(16), 10037–10072 (2021)
18. Bogdanos, C., Nesseris, S.: Genetic algorithms and supernovae type ia analysis. J. Cosmol. Astropart. Phys. **2009**(05), 006 (2009)
19. Brasil, C.R.S., Delbem, A.C.B., da Silva, F.L.B.: Multiobjective evolutionary algorithm with many tables for purely ab initio protein structure prediction. J. Comput. Chem. **34**(20), 1719–1734 (2013)
20. Cai, B., Pan, G.-L., Fu, F.: Prediction of the post-fire flexural capacity of rc beam using ga-bpnn machine learning. J Perform Const Facil (2020)
21. Cantó, J., Curiel, S., Martínez-Gómez, E.: A simple algorithm for optimization and model fitting: Aga (asexual genetic algorithm). Astron. Astrophys. **501**(3), 1259–1268 (2009)
22. Carr, B., Hart, W., Krasnogor, N., Hirst, J., Burke, E., Smith, J.: Alignment of protein structures with a memetic evolutionary algorithm. In: Proceedings of the 4th Annual Conference on Genetic and Evolutionary Computation, pp. 1027–1034 (2002)
23. Cassisi, S., Schlattl, H., Salaris, M., Weiss, A.: First full evolutionary computation of the helium flash-induced mixing in population II stars. Astrophys. J. **582**(1), L43 (2002)
24. Chan, C., Aleti, A., Heger, A., Smith-Miles, K.: Evolving stellar models to find the origins of our galaxy. In: Proceedings of the Genetic and Evolutionary Computation Conference, pp. 1129–1137 (2019)
25. Charbonneau, P.: Genetic algorithms in astronomy and astrophysics. Astrophys. J. Suppl. Ser. **101**, 309 (1995)
26. Chen, A., Choo, K., Astrakhantsev, N., Neupert, T.: Neural network evolution strategy for solving quantum sign structures. Phys. Rev. Res. **4**(2), L022026 (2022)
27. Chen, B., Wang, T., Li, C., Dai, H., Song, L.: Molecule optimization by explainable evolution. In: International Conference on Learning Representation (ICLR) (2021)
28. Chen, C., Ong, S.P.: A universal graph deep learning interatomic potential for the periodic table (2022). arXiv:2202.02450
29. Chen, J., Chen, Y., Xianchen, X., Zhou, W., Huang, G.: A physics-guided machine learning for multifunctional wave control in active metabeams. Extrem. Mech. Lett. **55**, 101827 (2022)
30. Chen, W., Hasanipanah, M., Rad, H.N., Armaghani, D.J., Tahir, M.M.: A new design of evolutionary hybrid optimization of svr model in predicting the blast-induced ground vibration. Eng. Comput. **37**
31. Cheng, G., Gong, X.-G., Yin, W.-J.: Crystal structure prediction by combining graph network and optimization algorithm. Nat. Commun. **13**(1), 1–8 (2022)
32. Comin, A., Hartschuh, A.: Efficient optimization of shg hotspot switching in plasmonic nanoantennas using phase-shaped laser pulses controlled by neural networks. Opt. Exp. **26**(26), 33678–33686 (2018)

33. de França, F.O., Aldeia, G.S.I.: Interaction-transformation evolutionary algorithm for symbolic regression. Evol. Comput. **29**(3), 367–390 (2021)
34. De Geyter, G., Baes, M., Fritz, J., Camps, P.: Fitskirt: genetic algorithms to automatically fit dusty galaxies with a monte carlo radiative transfer code. Astron. Astrophys. **550**, A74 (2013)
35. Demetriou, D., See, L., Stillwell, J.: A spatial genetic algorithm for automating land partitioning. International Journal of Geographical Information Science **27**(12), 2391–2409 (2013)
36. Demetriou, D., Stillwell, J., See, L.M.: LandParcelS: A Module for Automated Land Partitioning. University of Leeds, School of Geography (2012)
37. Deringer, V.L., Caro, M.A., Csányi, G.: Machine learning interatomic potentials as emerging tools for materials science. Adv. Mater. **31**(46), 1902765 (2019)
38. Desai, S., Strachan, A.: Parsimonious neural networks learn interpretable physical laws. Sci. Rep. **11**(1), 1–9 (2021)
39. Durrant, J.D., Amaro, R.E., McCammon, J.A.: Autogrow: a novel algorithm for protein inhibitor design. Chem. Biol. Drug Des. **73**(2), 168–178 (2009)
40. Dushatskiy, A., Alderliesten, T., Bosman, P.A.N.: A novel surrogate-assisted evolutionary algorithm applied to partition-based ensemble learning. In: Proceedings of the Genetic and Evolutionary Computation Conference, pp. 583–591 (2021)
41. Evans, B., Al-Sahaf, H., Xue, B., Zhang, M.: Evolutionary deep learning: a genetic programming approach to image classification. In: 2018 IEEE Congress on Evolutionary Computation (CEC), pp 1–6. IEEE (2018)
42. Faris, H., Ala'M, A.-Z., Heidari, A.A., Aljarah, I., Mafarja, M., Hassonah, M.A., Fujita, H.: An intelligent system for spam detection and identification of the most relevant features based on evolutionary random weight networks. Inf. Fus. **48**, 67–83 (2019)
43. Fishman, E.: Bikeshare: A review of recent literature. Trans. Rev. **36**(1), 92–113 (2016)
44. Gao, S., Song, S., Cheng, J., Todo, Y., Zhou, M.: Incorporation of solvent effect into multiobjective evolutionary algorithm for improved protein structure prediction. IEEE/ACM Trans. Comput. Biol. Bioinform. **15**(4), 1365–1378 (2017)
45. Gao, X., Lee, G.M.: Moment-based rental prediction for bicycle-sharing transportation systems using a hybrid genetic algorithm and machine learning. Comput. Ind. Eng. **128**, 60–69 (2019)
46. Genty, G., Salmela, L., Dudley, J.M., Brunner, D., Kokhanovskiy, A., Kobtsev, S., Turitsyn, S.K.: Machine learning and applications in ultrafast photonics. Nat. Photo. **15**(2), 91–101 (2021)
47. Ghimire, S., Deo, R.C., Downs, N.J., Raj, N.: Self-adaptive differential evolutionary extreme learning machines for long-term solar radiation prediction with remotely-sensed modis satellite and reanalysis atmospheric products in solar-rich cities. Remote Sens. Environ. **212**, 176–198 (2018)
48. Giuliani, M., Castelletti, A., Pianosi, F., Mason, E., Reed, P.M.: Curses, tradeoffs, and scalable management: Advancing evolutionary multiobjective direct policy search to improve water reservoir operations. J. Water Res. Plan. Manag. (2):04015050 (2016)
49. Guijo-Rubio, D., Durán-Rosal, A.M., Gutiérrez, P.A., Gómez-Orellana, A.M., Casanova-Mateo, C., Sanz-Justo, J., Salcedo-Sanz, S., Hervás-Martínez, C.: Evolutionary artificial neural networks for accurate solar radiation prediction. Energy **210**, 118374 (2020)
50. Guillén, A., Bueno, A., Carceller, J.M., Martínez-Velázquez, J.C., Rubio, G., Todero Peixoto, C.J., Sanchez-Lucas, P.: Deep learning techniques applied to the physics of extensive air showers. Astropart. Phys. **111**, 12–22 (2019)
51. Zahid Halim, Muhammad Nadeem Yousaf, Muhammad Waqas, Muhammad Sulaiman, Ghulam Abbas, Masroor Hussain, Iftekhar Ahmad, and Muhammad Hanif. An effective genetic algorithm-based feature selection method for intrusion detection systems. *Computers & Security*, 110:102448, 2021
52. Hegde, R.S.: Photonics inverse design: pairing deep neural networks with evolutionary algorithms. IEEE J. Sel. Top. Quantum Electron. **26**(1), 1–8 (2019)
53. Hosseini, S., Zade, B.M.H.: New hybrid method for attack detection using combination of evolutionary algorithms, SVM, and ANN. Comput. Netw. **173**, 107168 (2020)

54. Eduardo Raul Hruschka, Ricardo JGB Campello, Alex A Freitas, et al. A survey of evolutionary algorithms for clustering. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 39(2):133–155, 2009
55. Ya-Lan, H., Chen, L.: A nonlinear hybrid wind speed forecasting model using lstm network, hysteretic elm and differential evolution algorithm. Energy Conver. Manag. **173**, 123–142 (2018)
56. Huang, J., Zhou, M., Zhang, J., Ren, J., Vatin, N.I., Sabri, M.M.S.: The use of ga and pso in evaluating the shear strength of steel fiber reinforced concrete beams. KSCE J. Civil Eng. **26**(9), 3918–3931 (2022)
57. Huang, Y., Dong, W., Chen, M., Zhang, K., Fang, D.: Evolutionary optimization design of honeycomb metastructure with effective mechanical resistance and broadband microwave absorption. Carbon **177**, 79–89 (2021)
58. Ivezić, Ž., Elitzur, M.: Self-similarity and scaling behaviour of infrared emission from radiatively heated dust-i. theory. Mon. Not. R. Astronom. Soc. **287**(4), 799–811 (1997)
59. Joseph, T., Lazio, W., Cordes, J.M., Novak, J.: The genetic algorithm: searching for planets around pulsars. In: Bulletin of the American Astronomical Society, vol. 25, p. 1366 (1993)
60. Kang, S., Jeong, W., Hong, C., Hwang, S., Yoon, Y., Han, S.: Accelerated identification of equilibrium structures of multicomponent inorganic crystals using machine learning potentials. NPJ Comput. Mater. **8**(1), 1–10 (2022)
61. Khare, E., Yu, C.-H., Obeso, C.G., Milazzo, M., Kaplan, D.L., Buehler, M.J.: Discovering design principles of collagen molecular stability using a genetic algorithm, deep learning, and experimental validation. Proc. Nat. Acad. Sci. **119**(40), e2209524119 (2022)
62. Kılıç, F., Yılmaz, İ.H., Kaya, Ö.: Adaptive co-optimization of artificial neural networks using evolutionary algorithm for global radiation forecasting. Renew. Energy **171**, 176–190 (2021)
63. Kim, W., Baag, C.-E.: Rapid and accurate two-point ray tracing based on a quadratic equation of takeoff angle in layered media with constant or linearly varying velocity functions. Bull. Seismol. Soc. Am. **92**(6), 2251–2263 (2002)
64. Kim, W., Hahm, I.-K., Ahn, S.J., Lim, D.H.: Determining hypocentral parameters for local earthquakes in 1-d using a genetic algorithm. Geophys. J. Int. **166**(2), 590–600 (2006)
65. Young Dae Ko: An efficient integration of the genetic algorithm and the reinforcement learning for optimal deployment of the wireless charging electric tram system. Comput. Ind. Eng. **128**, 851–860 (2019)
66. Kondo, T.: Evolutionary design and behavior analysis of neuromodulatory neural networks for mobile robots control. Appl. Soft Comput. **7**(1), 189–202 (2007)
67. Kowalski, M., Rubin, D., Aldering, G., Agostinho, R.J., Amadon, A., Amanullah, R., Balland, C., Barbary, K., Blanc, G., Challis, P.J. et al.: Improved cosmological constraints from new, old, and combined supernova data sets. Astrophys. J. **686**(2), 749 (2008)
68. Kwon, Y., Kang, S., Choi, Y.-S., Kim, I.: Evolutionary design of molecules based on deep learning and a genetic algorithm. Sci. Rep. **11**(1), 1–11 (2021)
69. Le, T.C., Winkler, D.A.: Discovery and optimization of materials using evolutionary approaches. Chem. Rev. **116**(10), 6107–6132 (2016)
70. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**(7553), 436–444 (2015)
71. Lei, Z., Gao, S., Zhang, Z., Zhou, M.C., Cheng, J.: Mo4: a many-objective evolutionary algorithm for protein structure prediction. IEEE Trans. Evol. Comput. **26**(3), 417–430 (2021)
72. Li, W., Zhao, D., He, C., Andong, H., Zhang, K.: Advanced machine learning optimized by the genetic algorithm in ionospheric models using long-term multi-instrument observations. Remote Sens. **12**(5), 866 (2020)
73. Xia Li and Anthony Gar-On Yeh: Integration of genetic algorithms and GIS for optimal location search. Int. J. Geograph. Inf. Sci. **19**(5), 581–601 (2005)
74. Li, Y., Ooi, H.K., Tchagang, A.: Deep evolutionary learning for molecular design (2020). arXiv:2102.01011
75. Li, Y., Cheng, W., Yu, L.H., Rainer, R.: Genetic algorithm enhanced by machine learning in dynamic aperture optimization. Phys. Rev. Accel. Beams **21**(5), 054601 (2018)

76. Wanzenböck, R., Arrigoni, M., Bichelmaier, S., Buchner, F., Carrete, J., Madsen, G.K.H.: Neural-network-backed evolutionary search for srtio3(110) surface reconstructions. Digit. Disc. **8** (2022)
77. Marzouq, M., Bounoua, Z., Fadili, H.E., Mechaqrane, A., Zenkouar, K., Lakhliai, Z.: New daily global solar irradiation estimation model based on automatic selection of input parameters using evolutionary artificial neural networks. J. Clean. Produ. **209**, 1105–1118 (2019)
78. Mason, K., Duggan, J., Howley, E.: Forecasting energy demand, wind generation and carbon dioxide emissions in Ireland using evolutionary neural networks. Energy **155**, 705–720 (2018)
79. Metcalfe, T.S.: Genetic-algorithm-based light-curve optimization applied to observations of the w ursae majoris star bh cassiopeiae. Astronom. J. **117**(5), 2503 (1999)
80. Metcalfe, T.S., Charbonneau, P.: Stellar structure modeling using a parallel genetic algorithm for objective global optimization. J. Comput. Phys. **185**(1), 176–193 (2003)
81. Moayedi, H., Moatamediyan, A., Nguyen, H., Bui, X.-N., Bui, D.T., Rashid, A.S.A.: Prediction of ultimate bearing capacity through various novel evolutionary and neural network models. Eng. Comput. **36**(2), 671–687 (2020)
82. Mokiem, M.R., de Koter, A., Puls, J., Herrero, A., Najarro, F., Villamariz, M.R.: Spectral analysis of early-type stars using a genetic algorithm based fitting method. Astron. Astrophys. **441**(2), 711–733 (2005)
83. Moster, B.P., Naab, T., Lindström, M., O'Leary, J.A.: Galaxynet: connecting galaxies and dark matter haloes with deep neural networks and reinforcement learning in large volumes. Mon. Not. R. Astronom. Soc. **507**(2), 2115–2136 (2021)
84. Nanayakkara, V.K., Ikegami, Y., Uehara, H.: Evolutionary design of dynamic neural networks for evaporator control. Int. J. Ref. **25**(6), 813–826 (2002)
85. Neshat, M., Nezhad, M.M., Abbasnejad, E., Mirjalili, S., Tjernberg, L.B., Garcia, D.A., Alexander, B., Wagner, M.: A deep learning-based evolutionary model for short-term wind speed forecasting: a case study of the lillgrund offshore wind farm. Energy Convers. Manag. **236**, 114002 (2021)
86. Nesseris, S., Shafieloo, A.: A model-independent null test on the cosmological constant. Mon. Not. R. Astron. Soc. **408**(3), 1879–1885 (2010)
87. Nunkesser, R., Morell, O.: An evolutionary algorithm for robust regression. Comput. Stat. Data Anal. **54**(12), 3242–3248 (2010)
88. Owoyele, O., Pal, P., Torreira, A.V., Probst, D., Shaxted, M., Wilde, M., Senecal, P.K.: Application of an automated machine learning-genetic algorithm (automl-ga) coupled with computational fluid dynamics simulations for rapid engine design optimization. Int. J. Engine Res. **23**(9), 1586–1601 (2022)
89. Paton, F.L., Maier, H.R., Dandy, G.C.: Including adaptation and mitigation responses to climate change in a multiobjective evolutionary algorithm framework for urban water supply systems incorporating ghg emissions. Water Res. Res. **50**(8), 6285–6304 (2014)
90. Pegg, S.C.-H., Haresco, J.J., Kuntz, I.D.: A genetic algorithm for structure-based de novo design. J. Comput.-Aided Mol. Des. **15**(10), 911–933 (2001)
91. Podryabinkin, E.V., Tikhonov, E.V., Shapeev, A.V., Oganov, A.R.: Accelerating crystal structure prediction by machine-learning interatomic potentials with active learning. Phys. Rev. B **99**(6), 064114 (2019)
92. Puls, J., Urbaneja, M.A., Venero, R., Repolust, T., Springmann, U., Jokuthy, A., Mokiem, M.R.: Atmospheric NLTE-models for the spectroscopic analysis of blue stars with winds-II. line-blanketed models. Astron. Astrophys. **435**(2), 669–698 (2005)
93. Ren, H., Fojtik, M.: Standard cell routing with reinforcement learning and genetic algorithm in advanced technology nodes. In: Proceedings of the 26th Asia and South Pacific Design Automation Conference, pp. 684–689 (2021)
94. Rolnick, D., Donti, P.L., Kaack, L.H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A.S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A. et al.: Tackling climate change with machine learning. ACM Comput. Surv. (CSUR) **55**(2), 1–96 (2022)
95. Ruiz, A.N., Cora, S.A., Padilla, N.D., Domínguez, M.J., Vega-Martínez, C.A., Tecce, T.E., Orsi, Á., Yaryura, Y., Lambas, D.G., Gargiulo, I.D. et al.: Calibration of semi-analytic models of galaxy formation using particle swarm optimization. Astrophys. J. **801**(2), 139 (2015)

96. Ruuzek, B., Kvasnicka, M.: Differential evolution algorithm in the earthquake hypocenter location. Pure Appl. Geophys. **158**(4), 667–693 (2001)
97. Safaei, M., Rezayan, H., Firouzabadi, P.Z., Sadidi, J.: Optimization of species distribution models using a genetic algorithm for simulating climate change effects on zagros forests in Iran. Ecol. Inf. **63**, 101288 (2021)
98. Şahin, C.B., Dinler, Ö.B., Abualigah, L.: Prediction of software vulnerability based deep symbiotic genetic algorithms: phenotyping of dominant-features. Appl. Intell. **51**(11), 8271–8287 (2021)
99. Sahni, V., Starobinsky, A.: Reconstructing dark energy. Int. J. Mod. Phys. D **15**(12), 2105–2132 (2006)
100. Salazar, J.Z., Reed, P.M., Herman, J.D., Giuliani, M., Castelletti, A.: A diagnostic assessment of evolutionary algorithms for multi-objective surface water reservoir control. Adv. Water Res. **92**, 172–185 (2016)
101. Salcedo-Sanz, S., Cornejo-Bueno, L., Prieto, L., Paredes, D., García-Herrera, R.: Feature selection in machine learning prediction systems for renewable energy applications. Renew. Sust. Energy Rev. **90**, 728–741 (2018)
102. Sambridge, M., Gallagher, K.: Earthquake hypocenter location using genetic algorithms. Bull. Seismol. Soc. Am. **83**(5), 1467–1491 (1993)
103. Schlattl, H., Cassisi, S., Salaris, M., Weiss, A.: On the helium flash in low-mass population III red giant stars. Astrophys. J. **559**(2), 1082 (2001)
104. Seyedzadeh, S., Rahimian, Oliver, F.P.S., Rodriguez, S., Glesk, I.: Machine learning modelling for predicting non-domestic buildings energy performance: a model to support deep energy retrofit decision-making. Appl. Energy **279**, 115908 (2020)
105. Sonnenberg, J. et al.: Fundamentals of land consolidation as an instrument to abolish fragmentation of agricultural holdings (2002)
106. Soto, D., Soto, W.: Evolutionary algorithm for solving supervised classification problems: an experimental study. In: 2022 6th International Conference on Intelligent Systems, Metaheuristics and Swarm Intelligence, pp. 24–29 (2022)
107. Y. Sun, Y. Jiao, C. Shi, Y. Zhang, Deep learning-based molecular dynamics simulation for structure-based drug design against sars-cov-2. Comput. Struct. Biotechnol. J. (2022)
108. Szustakowski, J.D., Weng, Z.: Protein structure alignment using a genetic algorithm. Prot.: Struct. Funct. Bioinf. **38**(4), 428–440 (2000)
109. Tani, L., Rand, D., Veelken, C., Kadastik, M.: Evolutionary algorithms for hyperparameter optimization in machine learning for application in high energy physics. European Phys. J. C **81**(2), 1–9 (2021)
110. Theis, C.H., Kohle, S.: Multi-method-modeling of interacting galaxies-I. A unique scenario for NGC 4449? Astron. Astrophys. **370**(2), 365–383 (2001)
111. Ünal, H.T., Başçiftçi, F.: Evolutionary design of neural network architectures: a review of three decades of research. Artif. Intell. Rev. 1–80 (2021)
112. Wahde, M.: Determination of orbital parameters of interacting galaxies using a genetic algorithm-description of the method and application to artificial data. Astron. Astrophys. Suppl. Ser. **132**(3), 417–429 (1998)
113. Wahde, M., Donner, K.J.: Determination of the orbital parameters of the m 51 system using a genetic algorithm. Astron Astrophys. **379**(1), 115–124 (2001)
114. Wall, M.: Galib: a c++ library of genetic algorithm components. Mech. Eng. Dep. MA Inst. Technol. **87**, 54 (1996)
115. Wang, W., Tang, R., Li, C., Liu, P., Luo, L.: A BP neural network model optimized by mind evolutionary algorithm for predicting the ocean wave heights. Ocean Eng. **162**, 98–107 (2018)
116. Wang, Y., Chengru, W., Zhao, S., Wang, J., Bingfeng, Z., Han, M., Qing, D., Ni, M., Jiao, K.: Coupling deep learning and multi-objective genetic algorithms to achieve high performance and durability of direct internal reforming solid oxide fuel cell. Appl. Energy **315**, 119046 (2022)
117. Wierzbiński, M., Pławiak, P., Hammad, M., Acharya, U.R.: Development of accurate classification of heavenly bodies using novel machine learning techniques. Soft Comput. **25**(10), 7213–7228 (2021)

118. Wu, Z., Kan, S.B.J., Lewis, R.D., Wittmann, B.J., Arnold, F.H.: Machine learning-assisted directed protein evolution with combinatorial libraries. Proc. Nat. Acad. Sci. **116**(18), 8852–8858 (2019)
119. Hao, X., Chang, H., Zhang, D.: DLGA-PDE: Discovery of PDEs with incomplete candidate library via combination of deep learning and genetic algorithm. J. Comput. Phys. **418**, 109584 (2020)
120. Hao, X., Zhang, D.: Robust discovery of partial differential equations in complex situations. Phys. Rev. Res. **3**(3), 033270 (2021)
121. Hao, X., Zhang, D., Zeng, J.: Deep-learning of parametric partial differential equations from sparse and noisy data. Phys. Fluids **33**(3), 037132 (2021)
122. Haoran, X., Ma, J., Tan, P., Chen, B., Zhen, W., Zhang, Y., Wang, H., Xuan, J., Ni, M.: Towards online optimisation of solid oxide fuel cell performance: combining deep learning with multi-physics simulation. Energy AI **1**, 100003 (2020)
123. Yaseen, Z.M., Afan, H.A., Tran, M.-T.: Beam-column joint shear prediction using hybridized deep learning neural network with genetic algorithm. In: IOP Conference Series: Earth and Environmental Science, vol. 143, p. 012025. IOP Publishing (2018)
124. Yoon, E.J., Thorne, J.H., Park, C., Lee, D.K., Kim, K.S., Yoon, H., Seo, C., Lim, C.-H., Kim, H., Song, Y.-I.: Modeling spatial climate change landuse adaptation with multi-objective genetic algorithms to improve resilience for rice yield and species richness and to mitigate disaster risk. Environ. Res. Lett. **14**(2), 024001 (2019)
125. Yoshikawa, N., Terayama, K., Sumita, M., Homma, T., Oono, K., Tsuda, K.: Population-based de novo molecule generation, using grammatical evolution. Chem. Lett. **47**(11), 1431–1434 (2018)
126. Zhan, Z.-H., Li, J.-Y., Zhang, J.: Evolutionary deep learning: a survey. Neurocomputing **483**, 42–58 (2022)
127. Zhang, C.-X., Yuan, Y., Zhang, H.-W., Shuai, Y., Tan, H.-P.: Estimating stellar effective temperatures and detected angular parameters using stochastic particle swarm optimization. Res. Astron. Astrophys. **16**(9), 008 (2016)
128. Zhong, X., Velez, C., Acevedo, O.: Partial charges optimized by genetic algorithms for deep eutectic solvent simulations. J. Chem. Theory Comput. **17**(5), 3078–3087 (2021)
129. Zhou, Y., Zhou, N., Gong, L., Jiang, M.: Prediction of photovoltaic power output based on similar day analysis, genetic algorithm and extreme learning machine. Energy **204**, 117894 (2020)
130. Zuo, Y., Qin, M., Chen, C., Ye, W., Li, X., Luo, J., Ong, S.P.: Accelerating materials discovery with bayesian optimization and graph deep learning. Mater. Today **51**, 126–135 (2021)