Chapter 1

Introduction

Artificial Intelligence (AI) is now a buzzword in different domains that consist of different advanced learning techniques and has shown remarkable superiority in several real-world applications. Optimization and learning are two major paradigms of AI techniques used to address complex real-world problems, each with its own focus but frequently complemented by the other. These advanced techniques grouped under the umbrella of AI are shown in Figure 1.1. AI is now becoming a more effective and essential part of human life. Further, Machine Learning (ML) and deep learning have undoubtedly contributed to tremendous achievements in AI in recent years, and more contributions are likely to follow. They have demonstrated extraordinary superiority in various real-world applications like computer vision, medical diagnostic systems, agriculture, robotics, and many more.

On the other side, as technology advances, a vast amount of data is generated, raising the problem complexity and computational challenges of real-world applications. Many of these real-world applications now emerge as complex or large-scale optimization problems. Almost every AI-based application has optimization tasks in its core that need to be solved effectively and efficiently, as the overall performance and quality of the solution are highly influenced by the way they have been addressed. These

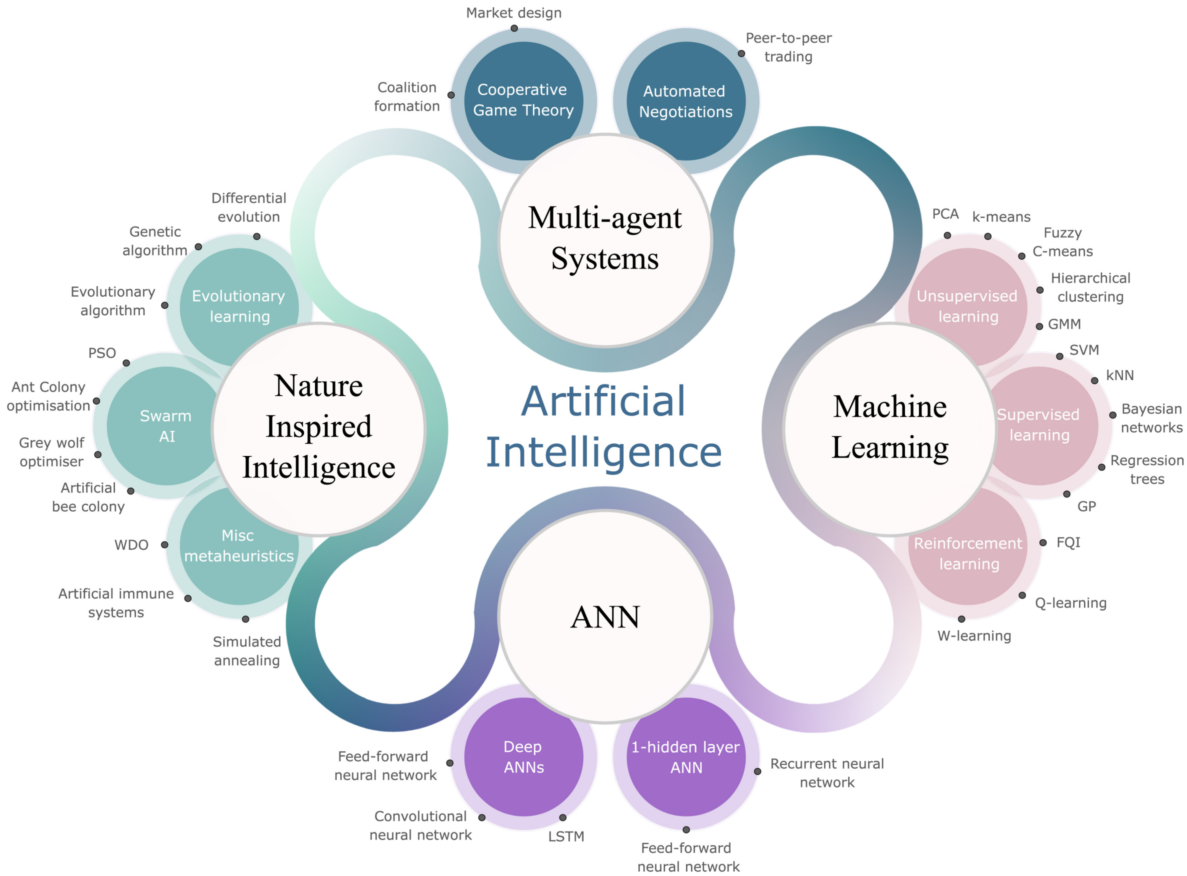

Figure 1.1: Groups of AI approaches [1]

optimization problems may vary from convex to non-convex, unimodal to multi-modal, single objective to multi-objective, small-scale problems to large-scale problems, and dynamic problems.

Among these problems, large-scale optimization problems have recently emerged as challenging in various application domains like logistic scheduling, trajectory optimization, pattern mining, and data security, as well as core optimization problems in AI such as structure learning of deep networks, feature selection, model parameter optimization, and many more. Numerous effective optimization techniques are already present to deal with convex optimization problems [2]. In contrast, non-convex problems and problems belonging to large-scale optimization tasks are still challenging and under the study of research communities for designing efficient approaches. Further, the emergence of Industry 4.0 gives a new edge to AI and the Internet of Things (IoT), which subsequently accelerates the development of multiple smart services that contribute to improving human life [3]. Advancements in AI and IoT also lead to the generation of huge volumes of data. Managing and analyzing these generated data is a challenging task. Thus, the revolution of Industry 4.0 opened new research avenues along with several challenges. Recently, it has drawn huge attention from research communities in optimization, deep learning, security, and many more [4].

Among these challenges, data privacy and its integrity are the major concern faced by smart digital services where data are communicated over a communication channel from different devices (sensors, IoT) and sources located in different geographical places. Generally, data privacy and its integrity have become more critical in e-healthcare systems. However, e-healthcare systems, Medi-cloud, and computer-aided diagnostic models were designed with the motivation to facilitate smart services to patients and medical experts by rendering the gap between doctors and patients across the globe for critical diagnosis suggestions. The major portion of the medical data is in the form of images which are being analyzed by AI techniques at either edge devices or clouds. Thus, exposing sensitive medical information to unauthorized access and manipulation of a single value may lead to a life threat to the patient due to an incorrect diagnosis. Further, hospitals, banks, electric power grid, and train systems are the country's critical national infrastructure which relies on the Supervisory Control and Data Acquisition (SCADA) and industrial control system to manage their production [5]. Subsequently, they are the primary target of cyberattacks, and recently, medical data thefts are growing widely [6], making data security and privacy one of the challenging areas in healthcare that needs to be addressed these days. Data encryption is one way to accomplish that, but these are suitable for text data; however, a large portion of healthcare data are in the form of images. Thus, making suitable changes and designing new approaches without losing information is also an open issue.

Another issue that has emerged as a result of the rise of big data is scalability [7].

Several areas, such as genomic data, are typically of high dimension and have a high potential for disease diagnosis. Currently, gene expression data analysis is a popular technique in health informatics [8, 9]. Due to the advancement of computational tools, the amount of such large-scale data has increased rapidly, making it difficult to use for analysis without suitable pre-processing. It leads to a curse of dimensionality issues for ML techniques and subsequently increases complexity and veracity problems besides other challenges such as storage, analysis, privacy, and security [10]. Further, only a small number of genes are sufficient biomarkers for specific disease identification. Similarly, in other domains, these voluminous data consist of noisy, redundant, and irrelevant features that adversely impact ML performance; thus, designing efficient techniques for analyzing large-scale data optimally are the prime research area.

Additionally, advanced ML models like convolutional networks, GAN, and Cyclic-GAN are the major attraction of the industry as well as academia due to their superior representation and learning ability. Among these, Cyclic-GAN has powerful potential to capture high-dimensional probability distribution leading to the generation of high perceptual quality images even for the cross-domain data [11]. However, it suffers from convergence pathologies resulting in unstable training [12]. To address this crucial problem manually is beyond the human scope as it is a type of large-scale optimization problem belonging to the NP-hard class. Most of these challenging problems can be solved as optimization problems efficiently, but due to the level of problem complexity and their characteristics, such as multiple objectives, multi-modal and non-convex loss functions, it is hard to solve using the traditional approach. They need some efficient and intelligent optimization techniques for solutions.

On the other hand, Evolutionary Computation (EC) methods are widely acknowledged as global optimization techniques which can be used for simple as well as complex optimization. ECs are inspired by natural phenomena and are broadly referred to as "Nature-Inspired Computations" (NIC) or "Nature-Inspired Algorithms" (NIA). The

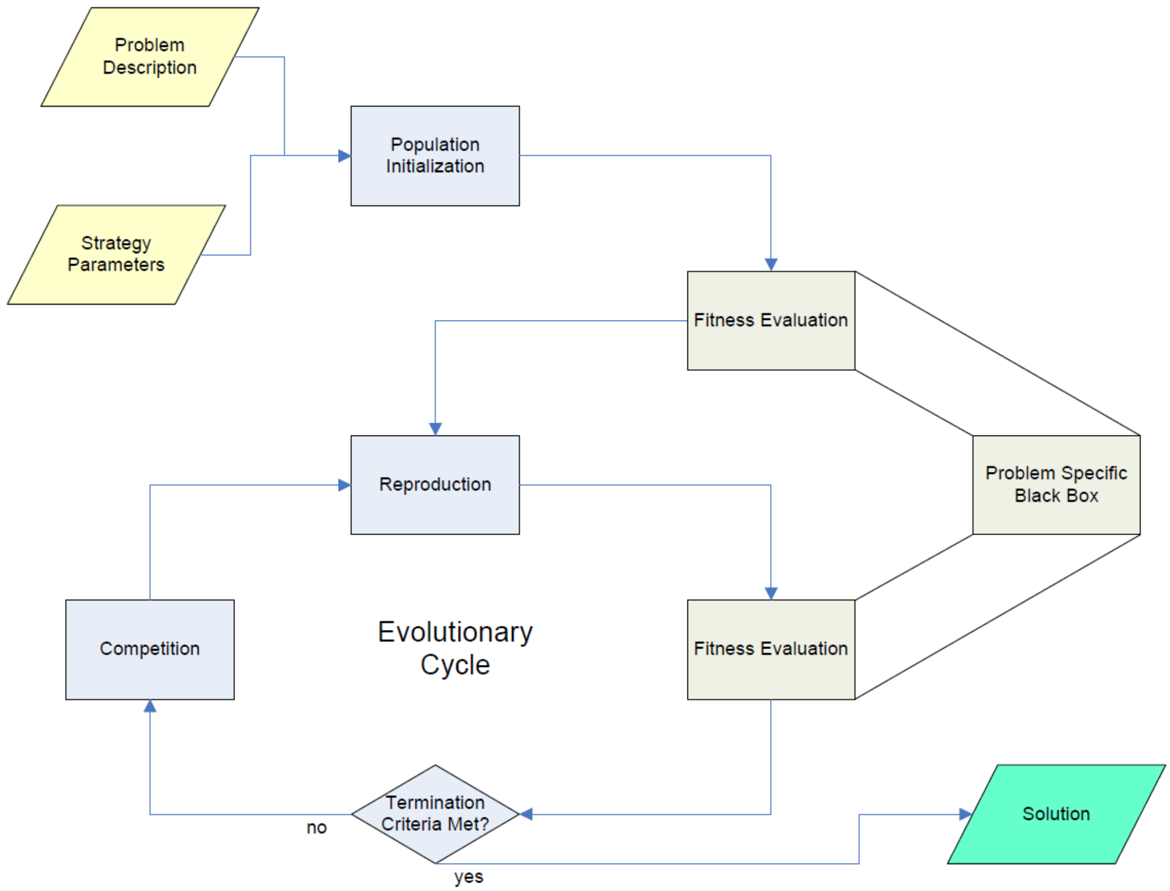

Figure 1.2: Overview of evolutionary problem-solving approach

basic overview of evolutionary problem solving is shown in Figure 1.2. Nowadays, they are widely applied to enhance the performance of ML tasks and have given rise to a new domain, "Evolutionary Machine Learning" [13].

EC methods are efficient in solving complex optimization tasks, but a few major challenges, such as slow convergence, local optima stagnation, and scalability issues, limit their broader applicability in advanced ML problems and their applications. Moreover, they are computationally expensive, which can be a hurdle in analyzing real-world problems related to big data, like high-dimensional medical data analysis or stream data analysis, etc. However, computation resources are now available to perform large computations. Since only utilizing high computation resources will not solve the whole purpose, we need to design efficient optimization techniques with good local search capabilities that are suitable, robust, and adaptive for modern applications, as well as that can utilize computation resources. Both ML and EC have their own strengths and limitations, which has sparked a surge in ongoing research in academia as well as industry to integrate these ideas to enhance their performance while overcoming limitations [14]. A few possibilities to deal with the challenges of EC are: (1) designing hybrids of existing ones by utilizing their strengths; (2) modern hybrids, which are the integration of interdisciplinary concepts like quantum computing from physics, chaos theory from mathematics, reinforcement learning, etc.; and (3) designing a new optimizer by taking inspiration from nature with good local search ability.

1.1 Need of Optimization

An optimization problem is a problem of finding the best possible solution while satisfying a set of constraints. Optimization problem which consists of only one objective function that needs to be minimized or maximized is known as a single-objective optimization problem. Mathematically, for a minimization problem, it is defined as:

$$\begin{aligned} & minimize_{\vec{x}}\, f(\vec{x}) \\ \text{subject to} \quad & \vec{x} \in X \subseteq \mathbb{R}^n \end{aligned} \tag{1.1}$$

and $\vec{x}^*$is an optimal solution in feasible set X for problem in (1.1) if it satisfies $f(\vec{x}^*) \leq f(\vec{x}) \quad \forall x \in X$. Numerous optimization techniques are available to solve the optimization problem presented in (1.1), ranging from the simplex method and classic gradient-based algorithms such as conjugate gradient, gradient descent, batch gradient descent, stochastic gradient descent, the Newton method, and the quasi-Newton method to EC approaches [15, 14].

The following points highlight the need for optimization in real-world applications:

1. Artificial intelligent techniques and the majority of real-world applications ranging

from engineering problems to big industrial data analysis have complex optimization problems within themselves that must be adequately addressed for better and more accurate analysis.

2. Most real-world problems naturally pose multi-objective optimization problems since they require the simultaneous satisfaction of several objectives.

3. A diverse range of domains, including feature selection, transportation, optimal key generation, portfolio optimization, supply chain management, agriculture, healthcare, industrial problems, complex networks such as social networks, wireless sensor networks, and many others, incentivize the research community to solve them as optimization problems. Hence optimization plays a vital role in almost every sector.

4. Efficient feature engineering simplifies the model by reducing training and execution times, input requirements, and computational costs. Furthermore, it not only improves the model's compactness and transparency by removing insignificant features from the dataset but also facilitates fast interpretation.

5. Because of their exceptional representation abilities, advanced ML models such as convolutional neural networks and generative models have led to significant breakthroughs in a wide range of domains. However, the network architecture's design is heavily reliant on the researcher's prior knowledge and expertise. Due to the constraint of inherent human knowledge, it is challenging for people to break free from the underlying thinking paradigm and build an optimal model.

6. Huge parameters in deep models must be optimally tuned, which is beyond the scope of human intelligence, necessitating the use of an automated scheme to address this issue.

1.1.1 Why EC over Other Traditional Optimization Techniques?

EC provides some advantages over traditional mathematical optimization techniques. A brief overview of the advantages is listed below, which presents a sufficient reason for utilizing EC to solve real-world optimization problems.

1. There is a wide range of high-performance optimization methods available for problems formulated as convex optimization [2]. However, many problems arising in important application domains pose non-convex problems. In particular, these problems are often noisy, stochastic and have limitations that change dynamically.

2. While some problems are in the form of single objective optimization, they are either non-convex or multi-modal in nature, and addressing such problems with traditional optimizers may result in suboptimal results.

3. ML is inherently a multi-objective task. However, traditionally, either only one of the objectives is adopted as the cost function, or multiple objectives are aggregated to a scalar cost function [16].

4. Traditional optimization techniques mainly rely on gradient-based information of the involved functions in order to find the optimal solution [17].

5. Traditional optimization algorithms are generally deterministic, and no randomness is used in generating new solutions, which can enhance the exploitation ability but lacks exploration.

6. Mathematical optimization approaches suffer from local optima entrapment. This refers to an algorithm assuming a local solution is a global solution, thus failing to obtain the global optimum. It is not effective for multi-modal non-convex problems.

7. Mathematical optimization methods are also often ineffective for problems with unknown or computationally expensive derivation.

8. The application diversity of EC algorithms is enormous, and the literature is growing quickly. EC has become a popular and powerful tool to deal with multi-objective optimizations (MOOs) due to their population-based search property, which makes an MOEA able to approximate the Pareto front in a single run [18].

9. This popularity is due to flexibility, gradient-free mechanism, and local optima avoidance of these algorithms [19], because they assume an optimization problem as a black box, and thus there is no need for derivative computation in search space. This quality makes them highly amenable to solving a diverse range of problems, such as highly nonlinear problems and tough optimization problems [19].

Thus, the primary focus of this thesis is to design efficient EC methods and validate their potential for a few challenging optimization problems in real-world and ML. The rest of the chapter is organized as follows. The next section presents the motivation of the thesis, followed by the challenges in Section 1.4 and the main objectives. Section 1.5 presents the contribution of this thesis, and Section 1.6 provides the thesis organization details.

1.2 Motivation

Several problems belonging to the NP-hard problem can be solved using EC methods where the classical derivative-based method fails to provide solutions. Nowadays, due to the emergence of advanced technology, voluminous data are generated, which need to be analyzed using ML techniques. It gives rise to a new field of research known as "Big Data analysis" and attracts worldwide attention from various domains. The recent pandemic has also accelerated research in this area for genomic data analysis. It also emphasizes the importance of automated tools that reduce the need for human intervention. Advanced ML models, on the other hand, such as convolutional neural networks, generative models, and other deep models, have demonstrated remarkable performance. However, these models need to tune numerous parameters to design an optimal and efficient model structure. They also involve multiple loss functions, which are non-convex and multi-modal in nature. Subsequently, it is hard to attain global optima, and analyzing the complexity of these models for better tuning is beyond the scope of humans and thus needs an intelligent mechanism to automate the whole process.

The rising trend is to combine the strength of EC into AI problems in order to overcome the underlying issue of the "curse of dimensionality," mode collapse, vanishing gradient, local optima stagnation in GAN, optimal key generation for data security, and so on. From the optimization perspective, many applications in the real world can be formulated as an optimization problems, which may vary from small-scale to large-scale optimization problems and can be solved by using EC methods. EC methods are phenomenal in solving complex and challenging problems. However, the limitations of premature convergence, local optima stagnation, and scalability issues open up new avenues for researchers belonging to the optimization field. Thus, the research communities have shown continuous efforts to improve the existing methods and explore new phenomena for designing novel optimizers. The development of new optimizers is also supported by NFL Theorem. Therefore, we are motivated to explore the potential of EC methods for large-scale problems and ML applications by designing efficient optimizers suitable for modern hi-tech applications.

The motivations of the present study can be summarized below:

1. **Large-scale optimization problem:** The advent of advanced technology brings enormous opportunities and challenges together. Big data analysis emphasizes the urgent need for scalable and efficient optimization techniques, as well as advanced ML models capable of solving large-scale problems. Furthermore, most

ML-based applications face the curse of dimensionality due to the presence of noisy, irrelevant, and redundant features in large-scale data. These noisy and redundant features adversely impact the predictive model's performance; hence selecting relevant feature subsets in high dimensions is merely impossible for humans. Besides that, advanced models developed for big data analysis comes with their own challenges of parameter tuning, which is typically a large-scale problem that necessitates an efficient technique to address such optimization.

2. **Trend of hybrid model:** Developing hybrid models to overcome the issues possessed by individuals seeks great attention and gives rise to new research domains such as Evolutionary AI, Evolutionary ML (EML), and similar efforts have been made within EC to overcome the issues of slow convergence, poor diversity, etc.

3. **Urgent need for modern hybrids:** Emergence of Industry 4.0 increase the level of problem complexity involved in real-world applications. To address the optimization problem inherited within these applications, optimization methods that are efficient, adaptive, scalable, stable, and reliable are needed. Recent trend shows that multidisciplinary concepts such as chaos theory, dynamic networks, reinforcement learning, cellular automata, parallel computing, and quantum computing concepts can be utilized to design hybrid models with improved performance.

4. **NFL in optimization :** NFL supports the development of novel optimizers. Exploration of natural phenomena around us needs to be explored for designing novel efficient optimizers which can solve single as well as multi-objective problems with good exploration-exploitation ability and are suitable for modern applications.

5. **Lack of paired data and realistic image generation:** Deep generative models are quite popular in different application domains due to their realistic image generation ability. It is an interesting model which is quite popular among

research communities. Usually, the availability of paired data in the case of old dead painting and rare novel diseases are minimal. Restoring old dead paintings and cross-domain image translation or unpaired image translation present enormous advantages and cost-effective solutions for different real-world applications. Generated images can be used for knowledge distillation without revealing the original data, and multi-modal medical imaging for accurate diagnosis are a few interesting applications that motivate us to choose this application. Further, Cyclic-GAN, widely used for unpaired image translation, suffers from training instability, mode collapse, and other issues which can be solved by using the potential of the EC approach.

1.3 Challenges

Both EC and ML have different characteristics. The primary challenges that these techniques face are summarised below.

1.3.1 Challenges with EC Methods

Despite the usefulness and popularity of EC methods, there are still significant challenges and bottlenecks with such algorithms, particularly from a theoretical standpoint. Though researchers understand the underlying processes of how such algorithms function in practice, it is unclear why they work and under what conditions they work. Furthermore, all EC techniques feature algorithm-dependent parameters whose values might impact the performance of the method under consideration. However, it is unclear what the ideal values or settings are, as well as how to modify these parameters for maximum performance and robustness. Figure 1.3 depicts the various approaches that could be taken to address the expensive evaluation in EC.

Apart from the challenges in the context of a theoretical gap, methods of EC also pose other vital challenges that limit their applicability to modern optimization prob-

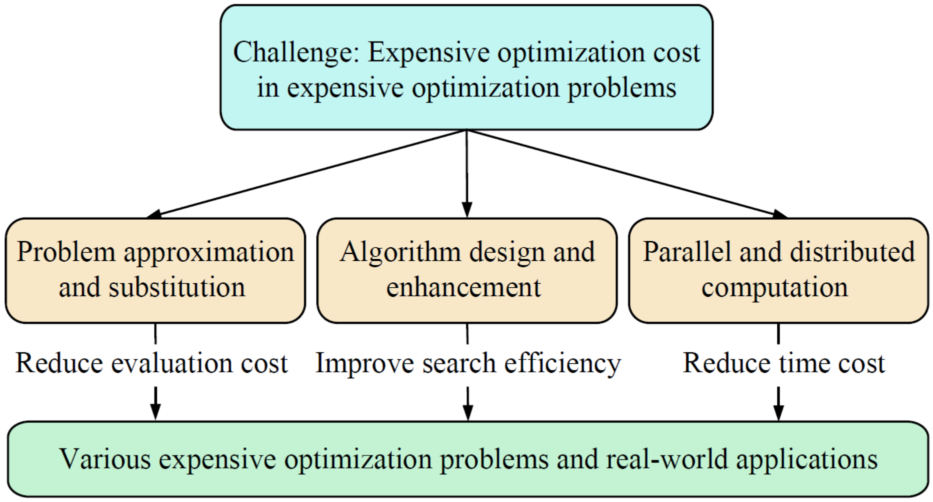

Figure 1.3: Directions for reducing high-cost optimization expenses in EC

lems and applications are as follows:

1. **Premature convergence and local optima stagnation:** These are common issues that become pernicious in the case of the multi-modal optimization task. Nowadays, numerous applications in engineering, AI, and other industrial problems involve non-convex multi-modal cost functions. It opens new avenues for researchers to study the search efficiency of EC methods, the role of diversity, and the balance between exploration-exploitation ability. Designing new intelligent search techniques or agent-assisted search mechanisms and exploring the hybridization of potential optimizers concepts are the recent trends in this direction.

2. **Computational expensive:** The computation cost is another serious concern of EC, mainly when dealing with feature selection problems and parameter optimization of deep networks because they involve many evaluations. Therefore, designing efficient and effective approaches to expensive tasks (feature selection, parameter optimization, security key generation) is still challenging, which can be solved by utilizing multidisciplinary concepts like parallel computing, quantum computing, and chaos theory. Designing a fast evaluation measure and efficient

search techniques are a few possibilities to address the expensive computation of EC approaches.

3. **Scalability:** The most prevalent concern is due to the emergence of Industry 4.0, which eases the generation of voluminous data with high dimensions. It also raises the complexity of the problems and becomes a more challenging task in ML, statistics, EC, and the biology community. Earlier in 1989, problems with more than fifteen dimensions were treated as large-scale problems [20]. Most of the existing EC methods were designed for small-scale problems and are not suitable for modern applications such as genomic data, where dimensions nearly reach thousand to million and thus become computationally expensive. Such problems demand an advanced search mechanism and efficient search space representation. Utilizing only the high and advanced computation resources will not solve this problem and thus necessitates the designing of novel algorithms and improvement in search mechanisms.

4. **Adaptive and reliable approach:** Few EC methods have great exploitation ability, but most are not adaptive and non-reliable. Thus limiting their applications to industrial large-scale problems and ML applications. During the search mechanism, each individual is not learning from the other experience and thus leads to the issue of local optima entrapment. Further, it is difficult to evaluate the whole process multiple times in large-scale problems. Providing a reliable, stable, and adaptive EC approach is another open problem that can be solved by exploring multidisciplinary concepts.

5. **Solution diversity and numerous tunable parameters:** Poor solution diversity is another problem, and accomplishing a good balance between diversification and intensification is essential to designing an efficient optimization technique. Besides, multiple tunable parameters in EC approaches without any theoretical

evidence on how to tune them also make them difficult for the researchers of the Non-EC domain.

1.3.2 Challenges in Real-world Large-scale Problems

This section explores the open challenges in a few demanding real-world problems and ML applications with optimization as a central theme.

1.3.2.1 Challenges with image privacy

Patients anticipate that healthcare organizations will abide by the rules to protect the privacy of their data and prevent its illegal use. In the modern era of health IoTs and cloud-based clinical diagnostic models, data privacy and security have thus become an open research area. Various encryption schemes are available for text data, which are not suitable for use directly on images due to their inherent characteristics. Moreover, in the healthcare domain, a majority of data is contributed in the form of images. Further, keys used for encryption are usually stored on the server, and a malicious server may lead to a data breach. As a result, securing images over smart digital services is the primary concern and difficult area to investigate from an optimization standpoint, where secure key plays a critical role. Searching for an optimal key for encryption is again a complex optimization that needs to be addressed effectively.

1.3.2.2 Challenges with feature selection

The revolution of Industry 4.0 and advancements in AI lead to the generation of huge volumes of large-scale data. Managing and analyzing this generated data is a challenging task. These data consist of redundant, noisy, and irrelevant features, which degrade the predictive performance of ML methods. Further, the presence of "complex interaction" between attributes makes this process difficult. The selection of two weakly relevant features, usually called complementary features, improves the predictive performance significantly. In contrast, selecting these two features may include redundant features, which subsequently increases the difficulty level of the problem. Apart from this, there are different approaches, like filter-based methods and wrapper-based approaches. The major limitations of the filter-based approach are that they do not consider the feature interaction property, and computing various metrics utilized in the filter-based method is much more complex for large-scale data due to the requirement of huge labeled instances. Wrapper-based approaches are proven to be more effective than other approaches. However, the high number of dimensions in large-scale data greatly worsens the behavior of EC approaches designed for a small-scale problem due to the exponential growth of search space. Overall, searching for subsets becomes expensive. Thus, there is a need for efficient optimization with exception search ability for large-scale problems.

1.3.2.3 Challenges with generative modeling

GAN [21] is a deep learning model that has emerged as one of the most promising approaches for unsupervised learning on complex distributions in the past few years. Through collaborative game learning of two modules in the framework: the generative model and discriminative model, i.e., generator and discriminator, the framework provides phenomenal outcomes. Deep neural networks are commonly employed as a generator or discriminators. The core notion behind GAN is that the generator and discriminator create a dynamic convex-concave game. Despite its widespread success in an extensive range of applications, GAN training remains a complex and demanding task because of the network's adversarial dynamics that can result in a variety of convergence pathologies [12], including mode collapse, vanishing gradient, and exploding gradients, causing instability [22]. Once gradient pathologies emerge, the generator is scarcely able to learn due to a lack of valuable gradient signal supplied by the discriminator, eventually leading to noise generation during the training procedures. These issues are typically caused by an imbalance between the discriminator and generator [23]. As a result, GAN training remains a difficult and open large-scale optimization problem, which becomes more challenging with Cyclic-GANs. Further, they fail in complex datasets of images with a salient object, such as medical images, because several current methods overlook the semantic information from the original image. Besides, model training requires high computation resources; thus, providing a low-cost solution in the recent computing world is another challenging area of research.

1.4 Objectives

The primary objectives of the study are listed as follows:

1. To investigate the concepts of hybridization by combining the individual strengths and unique characteristics of existing optimizers into a single optimization method to overcome existing limitations.

2. To examine the potential of multidisciplinary concepts such as Q-Learning, chaos theory, and quantum computing in conjunction with the EC method for the design of modern hybrids.

3. To propose novel evolutionary methods that are computationally efficient and suitable for both single-objective and multi-objective optimization problems and modern AI applications.

4. To validate the proposed methods on standard test functions, real-world datasets, and complex optimization tasks in real-world applications such as optimal key generation for image encryption and large-scale feature selection problems.

5. To propose a novel training scheme for Cyclic-GAN in the context of multi-objective optimization problems while addressing mode collapse, vanishing gradient, and training instability issues.

6. To propose a quantized version of the multi-objective Cyclic-GAN model for unpaired image translation and making it suitable for low-cost devices while using limited computation resources. Following that, investigate the impact of model quantization on generative performance and propose an intelligent scheme to compensate for the impact of quantization on overall performance.

7. To develop fast and efficient Non-dominated Sorting (focal step of MOEA) for multi-objective and many-objective optimization problems and modify it to accommodate cutting-edge technologies.

1.5 Thesis Contributions

All of the proposed methods have been adequately tested on standard benchmark public data and have shown improvements over the state-of-the-art methods. The major contributions of this study can be summarized as follows:

1. **Design and development of a novel GDWCN-PSO algorithm and its application to data security:** Premature convergence, local optima stagnation, and low diversity are EC methods' major and most common limitations. PSO is a popular EC approach among researchers from various application domains, which has led to the development of numerous PSO variants intending to improve its effectiveness. DWCN-PSO is one of these recent variants where particles were considered in the form of a dynamic network. However, it added the new characteristic of self-adaptation. Still, premature convergence limits its performance, which becomes pernicious in multi-modal optimization tasks where more than one global optima with several local optima may exist.

   Further, due to poor diversity, all particles converge at the same optima once any single particle starts converging. Following the recent trend of combining different approaches and operations to achieve greater efficiency, GDWCN-PSO has been proposed as a solution to the problem of premature convergence and local optimum stagnation in DWCN-PSO. GDWCN-PSO has been evaluated using standard benchmarks and real-world applications of optimal key generation. The application is chosen based on the importance and necessity of data privacy. Different medical images have been used to demonstrate the encryption process. A multi-objective variant of GDWCN-PSO is also proposed and tested using convergence and diversity performance metrics on standard test suits.

2. **Development of modern hybrids of Squirrel Search Algorithm by using modern interdisciplinary concepts for optimal feature subset selection:** The interdisciplinary concept Q-Learning (QL), a reinforcement learning component that can enhance the diversity of solutions, has been utilized to modify the local search of SSA. However, the cooperative nature of squirrels has been maintained similarly to the original SSA in the proposed adaptive QL-based SSA. Although the proposal QL-SSA offers stability, reliability, and adaptiveness and is suitable for decentralized problems. However, the computation time is noted to be high. Its efficacy was validated against 20 real-world benchmark datasets along with two large-scale genomic datasets for optimal feature selection problems. KNN and SVM (RBF) have been utilized as learning algorithms. Based on this observation, we further investigate the other multidisciplinary concept of chaos theory and quantum computing in conjunction with SSA. We incorporated chaos theory with SSA to overcome the premature convergence issue. Three chaotic maps have been investigated in the original SSA, which produced three chaotic versions of SSA (CSSA). Additionally, qubit representation and quantum gates have been integrated with SSA and CSSA individually, resulting in QSSA and QCSSA, respectively to enhance the population diversity ability and effective search capabilities. These modern hybrids are validated on large-scale genomic data for optimal feature selection problems. Experimental results demonstrate

the effectiveness of these methods for challenging feature selection problems.

3. **Design and development of novel efficient optimization algorithms by exploring the natural phenomena of bird migration:** A novel Murmuration-Flight based Dispersive Optimization (MDO) inspired by natural bird migration phenomena and flight patterns has been proposed for both single and multi-objective optimization. To the best of our knowledge, this work is the first of its kind to utilize *Lévy* flights to initialize the first population of solutions, thereby ensuring better exploration of the search space from the starting point of the optimization process. It has few tunable parameters compared to pre-existing optimizers, making it relatively simple to implement, understand, and suitable for the different application domains. Both proposals were investigated using standard benchmarks and test suits. A nonparametric test with 5% significance level has been performed to verify MDO. Further, the multi-objective version (MDO-M) is also tested against different performance metrics for convergence and diversity evaluation. MDO is also applied to optimal key generation for image encryption and optimal feature subset selection problems. A diverse set of datasets belonging to a different domain, including variable complexity in terms of the number of instances, number of classes, and feature dimensions, have been used for an optimal feature selection problem. The same problem is also solved using MDO-M by optimizing two objectives- the number of selected features and predictive model performance.

4. **Development of novel multi-objective Cyclic-GAN for unpaired image-to-image translation:** In this study, we addressed the open issues of Cyclic-GAN in the context of large-scale optimization problems. Training instability, local optima stagnation, vanishing gradient, mode collapse, and solution diversity issues are the primary concern with Cyclic-GAN. We employed EC techniques to address these concerns due to their potential to solve multi-objective optimization,

non-convex optimization problems, and many more. The whole training process is formulated as a multi-objective optimization problem. We introduced a new training approach by combining EC, multi-objective optimization, and Cyclic-GAN with different selection schemes resulting in the proposal of EMOCGAN. Two objectives, Frechet Inception Distance (FID) and Inception Score(IS), were optimized simultaneously using NSGA-II. Different losses are incorporated during training to generate more realistic images. EC method deals with training instability and mode collapse problem. Further, metropolis acceptance criteria with Pareto-based selection helped to overcome the local optima stagnation problem. EMOCGAN was verified on unpaired image-to-image translation problems by considering three popular real-world image datasets: Apple to Orange, Summer to Winter, and Monet to Picture.

Further, we have extended this model by incorporating three objective functions to enhance the model's generating ability. Because of the current industrial demand for IoT-based applications and the popularity gained by generative models, a quantized version of multi-objective Cyclic-GAN is also designed. To overcome the performance loss due to quantization, an intelligent selection mechanism that takes account of the gradient information has been introduced. This quantized version is tested on the same real-world datasets. It is also tested for three other medical datasets for unpaired image-to-image translation. The proposals effectively retain texture, salient objects, background, and color information and lead to more realistic image generation. It can be generalized to other tasks.

5. **Development of CUDA accelerated non-dominated sorting:** CUDA accelerated parallel corner sort method has been designed and developed to enhance the computation efficiency of the corner sort algorithm. Usually, non-dominated sorting plays a crucial role in multi-objective and many-objective optimization. Thus, due to enormous data generation in industrial problems posing many-objective

optimization at its core, we have studied and found the scope of parallelism in corner sort and employed CUDA to accelerate operations. The method was tested on a randomly generated large dataset. A theoretical study has been performed for complexity analysis of the technique. However, due to limited data availability, its efficiency must be verified on large industrial data sets, and it is planned to expand and validate on the same.

1.6 Organization of the Thesis

This thesis has been organized into the following eight chapters:

Chapter 1 provides the overall introduction to this thesis consisting of EC and its relevance in large-scale optimization and ML, followed by motivations, challenges, research objectives, and a summary of significant contributions.

Chapter 2 provides a theoretical background comprising basic definitions along with a comprehensive overview of EC techniques and deep learning. It also elucidates a detailed literature survey focusing on the EC techniques for large-scale optimization and ML problems. The comparative analysis of these approaches helps in identifying the research gap, presented in a summary section.

Chapter 3 introduces the GDWCN-PSO algorithm for single and multi-objective optimization problems and investigates it using standard benchmarks. Considering the premature convergence issue of PSO and its variant, DWCN-PSO, which is very detrimental in multi-modal optimization problems, the concept of Genetic Algorithm has been integrated with DWCN-PSO after each update to improve convergence with diversity. In order to demonstrate the applicability of GDWCN-PSO in real-world applications, it is further applied to solve the optimal key generation for medical image encryption.

Chapter 4 focuses on designing more stable, robust, self-adaptive, and scalable optimizers that are suitable for large-scale optimization problems in ML. In the context of big data problems where data dimensions may exceed thousands, the repetitive formation of a dynamic particle's network may become computationally expensive. Thus, in view of recent development, the SSA has better convergence ability than most of the pre-existing methods, making it more suitable for further improvement and modern applications. In order to attain stability, self-adaptation, scalability, and exceptional convergence ability with good diversity, different potential interdisciplinary concepts such as QL, chaos theory, and quantum computing have been integrated with SSA, resulting in new modern hybrids- QL-SSA, CSSA, QSSA, and QCSSA, respectively. These methods are briefly discussed in the context of optimal feature subset selections. Real-world benchmark datasets for ML and large-scale genomic datasets are utilized for this purpose.

Chapter 5 highlights the need for a novel and efficient optimization method. In contrast to the methods proposed in Chapter 4, which have the capability of achieving the desired objective needed for a challenging large-scale problem, they also increase the conceptual complexity that is natural in hybrid models. People apart from core optimization and ML involved in applied research need efficient optimizers with desirable characteristics for low-end devices while maintaining conceptual simplicity. Thus, in a similar spirit, this chapter introduces novel and efficient optimization approaches for both single-objective and multi-objective problems by modeling the natural phenomena of bird migration and flight patterns among birds. They are extensively investigated on single-objective benchmarks and multi-objective test suits. Performance of the multi-objective variant is extensively measured on different performance metrics. Further, the application of optimal-key generation for image encryption and optimal feature selection were presented. In-depth analyses were presented to validate their efficacy for modern applications.

Chapter 6 addresses another more challenging large-scale optimization problem to validate the efficiency of EC. Despite the success of generative models, training instability, mode collapse, and vanishing gradient are the major open issue. EC methods have great potential to solve problems with unknown solutions. They are suitable for such large-scale and non-convex problems belonging to the NP-hard class. This chapter designed a new training scheme with a different selection scheme and incorporated additional loss functions in Cyclic-GAN for unpaired image translations. The whole process has been designed and discussed as a multi-objective optimization problem. In the later part of the chapter, a quantized version with a new intelligent mechanism has been discussed while considering three different objective functions for unpaired image translation suitable for IoT devices.

Chapter 7 highlights the importance of parallel algorithms for many-objective optimization problems. Non-dominated sorting is a focal step as well as a computationally expensive part of multi-objective and many-objective optimization algorithms. In order to make it efficient, advanced and modern technology CUDA has been utilized to reduce the computation overhead of the corner sort algorithm (type of nondominated sorting) and introduced CUDA accelerated corner sort algorithm. Due to limited data availability, its efficiency is still needed to verify on large industrial data and is planned to be extended and validated on the same.

Chapter 8 summarises the thesis work with promising research directions in the area of EC and large-scale optimization problems in AI.