Metaheuristic Algorithms with Applications

www.maa.reapress.com

Metaheur. Algor. Appl. Vol. 2, No. 1 (2025) 99–111.

Paper Type: Original Article

Feature Selection with Metaheuristic Algorithms: A Review of Recent Developments (2020–2025)

Saikat Gochhait*

Federal State Budgetary Educational Institution of Higher Education, Samara State Medical University, Ministry of Healthcare, Samara, Russia; saikat.gochhait@sidtm.edu.in.

Citation:

Received: 20 June 2024 *Revised: 25 August 2024* *Accepted: 18 September 2024*	Gochhait, S.(2025). Feature selection with metaheuristic algorithms: a review of recent developments (2020–2025). *Metaheuristic Algorithms with Applications*, 2(1), 99-111.

Abstract

Feature selection is a critical preprocessing step in machine learning, aimed at identifying relevant features from high-dimensional datasets to improve model performance and reduce computational cost. Due to its Nondeterministic Polynomial (NP)-hard nature, metaheuristic algorithms have gained prominence for efficiently navigating the vast search space. This review examines approximately 150 metaheuristic algorithms developed or refined between 2020 and 2025, categorized into evolutionary, physics-based, human-social, and swarm intelligence approaches. Swarm intelligence algorithms dominate recent advances, comprising 55% of the surveyed methods, reflecting their scalability and effectiveness in complex domains such as healthcare and cybersecurity. The review highlights algorithmic trends, including hybridization, chaos-based diversity enhancement, and multi-objective optimization, and proposes future directions focused on adaptive, interpretable, and Artificial Intelligence (AI)-integrated frameworks.

Keywords: Metaheuristic, Optimization, Feature selection, Nondeterministic polynomial-hard, Machine learning.

1 | Introduction

The rapid evolution of Artificial Intelligence (AI), deep learning, and big data analytics has transformed modern science and industry, leading to an era defined by immense data complexity. In fields such as genomics, medical imaging, cybersecurity, and financial forecasting, datasets now contain thousands of feature attributes, signals, or measurements that collectively describe a phenomenon or system. While this abundance of information offers opportunities for powerful predictive modeling, it also introduces severe computational and analytical challenges. High-dimensional data often contain redundant, irrelevant, or noisy features that degrade model accuracy, increase training time, and complicate interpretability. As models grow deeper and data become larger, the need for efficient dimensionality reduction techniques becomes increasingly vital [1].

Corresponding Author: saikat.gochhait@sidtm.edu.in
https://doi.org/10.48313/maa.v2i1.34

Feature selection has emerged as one of the most effective solutions to this problem. It seeks to identify a subset of the most informative features while discarding those that contribute little or no value to prediction or classification. Unlike feature extraction, which transforms data into new dimensions, feature selection preserves the original meaning of each variable, maintaining interpretability. By eliminating redundancy and focusing on relevance, feature selection improves learning efficiency, enhances generalization, and reduces overfitting, key requirements for dependable and explainable machine learning. This makes it indispensable in domains where decisions must be both accurate and transparent, such as medical diagnostics, credit scoring, industrial control, and environmental modeling [2].

However, identifying the optimal subset of features remains one of the most challenging problems in data analysis. The search space of possible feature combinations grows exponentially with the number of available features, making exhaustive enumeration computationally infeasible even for moderately sized datasets. This challenge belongs to the class of Nondeterministic Polynomial (NP)-hard problems, meaning that no known deterministic algorithm can guarantee an optimal solution within polynomial time. Traditional methods such as forward selection, backward elimination, or branch-and-bound approaches often fail to scale to modern data volumes. Consequently, researchers have turned to metaheuristic algorithms and stochastic optimization methods that efficiently approximate near-optimal solutions in vast, complex, and nonlinear spaces.

Metaheuristic algorithms have gained prominence in feature selection because they rely on intelligent sampling and adaptive search strategies rather than exhaustive enumeration. They operate through iterative exploration of candidate solutions, guided by randomized movements and feedback from the objective function. These algorithms are inherently flexible and require little prior knowledge of the problem structure. They balance two critical behaviors: exploration, which encourages diversity and prevents premature convergence to local optima, and exploitation, which intensifies search around promising regions. Achieving an effective equilibrium between these two processes is central to the success of any metaheuristic method [3], [4].

From 2020 to 2025, research in metaheuristic optimization has expanded dramatically. Hundreds of new algorithms have been introduced, inspired by mechanisms observed in nature, human society, and physical systems. This period represents a turning point in the evolution of metaheuristics, characterized by a shift from static, rule-based models toward dynamic, hybrid, and intelligent frameworks. The landscape now extends far beyond the classical algorithms of the past, such as genetic algorithms, particle swarm optimization, and ant colony optimization. Contemporary designs integrate concepts from behavioral psychology, swarm intelligence, physics, medicine, and even social interaction, creating a diverse taxonomy of optimization strategies [5].

These algorithms can be broadly grouped into four major categories. Evolutionary algorithms simulate the processes of biological evolution, using operators analogous to mutation, selection, and recombination to evolve solutions across generations. Physics-based algorithms draw on natural laws such as gravity, equilibrium, and thermodynamics to guide particle movement and energy transfer within a search space. Human–social algorithms imitate collective human behavior, learning, and cooperation, often incorporating decision-making, negotiation, or educational dynamics. Finally, swarm intelligence algorithms mimic the self-organized movement of social animals such as birds, fish, and insects, emphasizing distributed problem-solving and communication among agents [6].

Each of these categories contributes a unique perspective on optimization. Evolutionary algorithms excel at maintaining diversity through genetic operators, making them effective for multimodal and non-convex problems. Physics-based algorithms introduce deterministic mathematical formulations that ensure stable convergence and predictable dynamics. Human–social algorithms leverage cognitive and cooperative processes that adapt well to dynamic or imbalanced environments, such as data streams or evolving networks. Swarm intelligence algorithms, in contrast, are particularly suited for large-scale problems, as their decentralized nature promotes scalability and parallelism. Together, these families form a comprehensive toolkit for tackling the feature selection problem in both continuous and discrete domains [7].

Feature selection using metaheuristics has gained traction because these algorithms are naturally well-suited to the combinatorial nature of subset selection. Each feature subset can be represented as a binary vector, where each bit indicates whether a feature is included or excluded. The optimization objective is typically twofold: maximizing classification accuracy while minimizing the number of selected features. This dual objective can be addressed either by scalarizing both objectives into a single weighted function or by adopting multi-objective optimization frameworks that yield Pareto-optimal solutions. The latter approach has become increasingly popular because it offers greater flexibility and deeper insight into the trade-offs between accuracy and complexity [8].

The period from 2020 to 2025 has seen several innovations in the design and application of metaheuristics for feature selection. One major advancement has been the incorporation of chaotic systems and random perturbations to enhance population diversity and prevent stagnation. Chaos theory introduces deterministic randomness: small changes in initialization lead to large differences in trajectories, helping the algorithm escape local minima. Another significant direction has been the development of hybrid algorithms, combining two or more metaheuristics to exploit complementary strengths. For example, hybrid models may use swarm-based exploration to maintain diversity and evolutionary refinement to achieve convergence. Such combinations have yielded substantial improvements in accuracy and speed, especially for high-dimensional datasets [1].

Adaptive and self-regulating algorithms have also become a defining trend of the 2020–2025 period. Unlike static algorithms with fixed control parameters, adaptive metaheuristics adjust their parameters dynamically during the search process, guided by performance feedback or environmental cues. Reinforcement learning has been integrated into several designs to enable the algorithm to learn from its previous actions, improving operator selection and convergence strategy. This convergence between machine learning and optimization has created a new class of intelligent metaheuristics capable of self-improvement and cross-domain adaptability [8].

Swarm-based algorithms dominate recent research, accounting for more than half of newly proposed metaheuristics in the past five years. They are particularly attractive for feature selection because of their population diversity, distributed computation, and ease of binary adaptation. Popular examples include the Sparrow Search Algorithm, Marine Predators Algorithm, and Slime Mould Algorithm, each offering distinct exploration mechanisms modeled after animal behavior. Variants of these algorithms often incorporate adaptive inertia weights, chaotic maps, or opposition-based learning to improve performance on feature selection benchmarks. Their scalability has made them the preferred choice for applications in gene selection, medical image analysis, and cybersecurity, where high-dimensional datasets and computational efficiency are critical [9].

Human–social algorithms, on the other hand, have expanded the conceptual reach of metaheuristics beyond biological analogies. They replicate processes such as political negotiation, educational competition, and human memory retention. This class of algorithms has shown promise in handling data imbalance and dynamic decision spaces, as their structures naturally integrate elements of learning and feedback. They often introduce memory archives, role-based cooperation, or competition-driven selection, making them adaptive and interpretable, two qualities increasingly valued in machine learning applications [10].

Evolutionary and physics-based algorithms remain highly influential as well. Evolutionary models have evolved, incorporating psychological and sociological elements to model motivation, cooperation, and emotion, resulting in richer dynamics. Physics-based algorithms continue to offer mathematically elegant frameworks for simulating energy balance, gravitational fields, and molecular interactions to guide search trajectories. Binary and discrete adaptations of these algorithms have made them especially useful for feature selection, where decisions are fundamentally categorical.

Another defining characteristic of this era is the rise of hybrid and ensemble frameworks. Researchers increasingly integrate multiple algorithms or learning models into a single system to enhance robustness.

Hybridization allows one algorithm to compensate for another's weaknesses. For example, an evolutionary component may maintain global diversity, while a swarm component accelerates convergence. In some advanced implementations, metaheuristics are coupled with deep learning architectures, where optimization guides the selection of input features or hyperparameters for neural networks. This merging of optimization and learning blurs the boundary between searching and training, creating adaptive frameworks that evolve over time [11], [12].

Interpretability has become a parallel objective alongside accuracy and efficiency. In many real-world scenarios, particularly in medicine and finance, understanding why a feature subset is chosen is as important as achieving high predictive accuracy. Metaheuristic-based feature selection offers a natural path toward explainability, as it provides explicit binary decisions on which features to retain or remove. When combined with statistical or visualization tools, these results can reveal meaningful patterns, such as identifying biomarkers for disease or indicators for financial risk. The transparency of this process strengthens the trustworthiness of AI models and facilitates regulatory compliance in sensitive applications [13].

The applications of metaheuristic feature selection between 2020 and 2025 have been diverse and impactful. In bioinformatics, these methods have been used to select minimal gene subsets that preserve diagnostic accuracy, enabling cost-effective and interpretable disease prediction. In medical imaging, they reduce the dimensionality of texture or wavelet features extracted from MRI and CT scans, improving the efficiency of classification models. In cybersecurity, adaptive swarm-based algorithms optimize subsets of network traffic attributes for intrusion detection, achieving high detection rates with minimal computational overhead. In industrial engineering, human–social and physics-based algorithms support process optimization and fault detection by identifying the most relevant monitoring variables. These applications highlight the versatility of metaheuristics and their capacity to generalize across domains with varying data structures [14].

Despite these advances, challenges remain. Many metaheuristic algorithms still rely heavily on random initialization, making their performance sensitive to initial conditions. Parameter tuning can be computationally expensive, and convergence behavior can vary significantly across problems. Scalability to extremely high-dimensional data, beyond tens of thousands of features, remains an open issue. Moreover, while hybrid and adaptive algorithms show promise, their increased complexity can make them difficult to analyze theoretically and to reproduce experimentally. Future research is expected to focus on improving stability, reducing energy consumption in computation, and establishing standardized evaluation frameworks to ensure comparability across studies [15].

Looking ahead, several emerging directions define the future of metaheuristic-based feature selection. The first is the integration of self-supervised and transfer learning mechanisms, enabling algorithms to reuse knowledge across similar problems or domains. The second is the incorporation of explainable AI principles, ensuring that optimization outcomes are interpretable and justifiable. The third involves developing lightweight, energy-efficient versions of metaheuristics suitable for real-time and embedded applications. Finally, the coupling of metaheuristics with quantum computing and neuromorphic architectures represents a frontier where optimization can leverage the next generation of computational paradigms.

In summary, feature selection remains a cornerstone of modern data-driven modeling, serving as an essential bridge between raw data and effective learning. The progress achieved from 2020 to 2025 in metaheuristic optimization has redefined how feature subsets are discovered and evaluated. The shift toward hybrid, adaptive, and interpretable frameworks marks the beginning of a new generation of intelligent search system algorithms that do not merely explore data but learn, reason, and adapt within it. Through these developments, metaheuristics are evolving from problem-solving tools into cognitive frameworks that embody the very principles of AI itself.

The remainder of this paper is organized as follows. Section 2 reviews the most relevant works on metaheuristic algorithms and their applications to feature selection, providing a comparative overview of recent trends from 2020 to 2025. Section 3 introduces the prominent families of metaheuristic algorithms, including evolutionary algorithms, physician-inspired algorithms, human and social-based algorithms, and swarm intelligence algorithms, each discussed with representative methods and recent advancements. Section 4 presents a critical discussion and outlines the advantages, limitations, and emerging research directions of these algorithms in feature selection. Finally, Section 5 concludes the paper and highlights future perspectives in developing hybrid and intelligent optimization strategies for high-dimensional data problems.

2 | Literature

In recent years, the explosive growth of high-dimensional data has intensified the demand for efficient feature selection and optimization algorithms. From 2020 to 2025, the research landscape has evolved toward integrating machine learning with intelligent optimization methods to enhance performance, interpretability, and scalability. Traditional statistical or deterministic approaches have increasingly proven insufficient for handling complex, nonlinear, and noisy datasets common in biomedical imaging, genomics, and text mining. Consequently, metaheuristic optimization has become a dominant paradigm bridging search efficiency with adaptability across domains [5].

The literature during this period reveals two converging research paths: first, the refinement of feature selection methods that reduce dimensionality while preserving discriminative power; and second, the development of metaheuristic optimization algorithms that balance exploration and exploitation across vast solution spaces. Together, these directions reflect a shift from isolated algorithmic design toward hybrid, adaptive, and interpretable frameworks, a hallmark of post-2020 machine learning research [1].

2.1 | Feature Selection

Feature selection is a core preprocessing technique in machine learning, identifying key attributes for tasks like classification and regression while eliminating irrelevant or redundant ones to reduce dimensionality and costs. Between 2020 and 2025, it has been crucial for handling complex datasets in genomics and imaging.

Methods include supervised (label-based relevance), unsupervised (structure-based), and semi-supervised. Filter methods use statistics; wrapper methods evaluate subsets with models; embedded methods integrate selection into training. Metaheuristics have become prominent for search strategies that avoid local optima in combinatorial spaces. Challenges such as imbalance and scalability have been addressed through hybrid approaches. In biomedicine, it identifies biomarkers; in text analysis, it improves sentiment extraction. The period reflects integration with deep learning for efficiency and interpretability [8], [16].

2.2 | Optimization Algorithms

Optimization ranges from polynomial to NP-hard problems, where metaheuristics excel by balancing exploration and exploitation. From 2020 to 2025, advancements include adaptive and hybrid designs in line with the no-free-lunch theorem.

Categories

Evolutionary (genetic evolution), physics-based (physical laws), human-social (human interactions), and swarm intelligence (animal collectives). They support continuous-to-discrete applications, with multi-objective variants for feature selection trade-offs, achieving superior benchmark performance. *Fig. 1* is a flowchart of metaheuristic categories, branching into Evolutionary (mutation/selection), physics-based (laws such as gravity), human-social (learning/competition), and swarm intelligence (flocking/foraging). Arrows show hybrid possibilities, aiding algorithm choice for feature selection [14], [17].

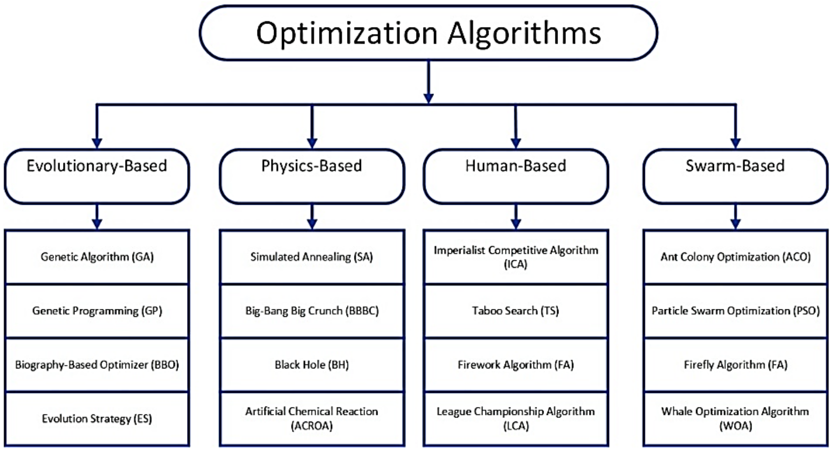

Fig. 1. Lowchart of metaheuristic algorithm categories and their hybrid interrelations.

3 | Optimization Algorithms (Focused on 2020–2025 Developments)

Post-2020 metaheuristics incorporate chaos and hybridization for feature selection. They start with random solutions, using exploration to explore new areas and exploitation to refine them. Inspirations include nature, physics, human behavior, and swarms. Evolutionary use of mutation-selection with chaos to avoid premature convergence, physics-based model energy-forces. Human-Social integrates memory and social cooperation. Swarm capitalizes on agent interactions to build collective intelligence. This diversity supports feature selection tasks [1].

4 | Optimizations on Feature Selection

This section reviews metaheuristic algorithms proposed between 2020 and 2025 for feature selection, organized by algorithmic category. Each method is discussed in terms of its mechanism, representation scheme, and empirical performance. Most algorithms employ binary or Boolean encodings to represent candidate feature subsets, where the optimization process seeks to maximize classification accuracy while minimizing the number of selected features [1], [8].

4.1 | Evolutionary Methods

Evolutionary methods evolve solutions via biological processes. Post-2020, they include hybridization and chaos for high-dimensional feature selection. The student psychology-based optimization algorithm, published in 2020, is inspired by students' psychological learning. It initializes populations as feature subsets, evolves through learning phases, evaluates fitness using classifiers, and adapts to achieve convergence. In feature selection, it optimizes classification by mimicking student motivation and is applied to engineering tasks to produce reduced subsets. The cooperation search algorithm, from 2021, is inspired by cooperative behaviors. It initializes cooperative groups as subsets, applies collaborative search, evaluates, and refines. Used in numerical optimization, it supports feature selection in large-scale problems. The love evolution algorithm, from 2024, is inspired by human relationships. It initializes solutions, evolves through stimulus-value-role operators, evaluates, and optimizes. In feature selection, it balances multiple objectives for engineering applications [18].

Table 1. lists these evolutionary algorithms with names, abbreviations, and years.

Algorithm Name	Abbreviation	Publication Year
Student psychology-based optimization algorithm	SPOA	2020
Adolescent identity search algorithm	AISA	2020
TDSD: triple distinct search dynamics	TDSD	2020
Red deer algorithm	RDA	2020
Cooperation search algorithm	CSA	2021
Linear prediction evolution algorithm	LPE	2021
Rat swarm optimizer	RSO	2021
Dingo optimizer	DO	2021
Artificial lizard search optimization	ALSO	2021
Artificial gorilla troops optimizer	AGT	2021
Battle Royale Optimization Algorithm	BRO	2021
Elephant clan optimization	ECO	2021
Golden eagle optimizer	GEO	2021
Horse herd optimization algorithm	HHO	2021
Ebola optimization search algorithm	EOSA	2022
Artificial yellow ground squirrel	YGSA	2022
Seasons optimization algorithm	SOA	2022
American zebra optimization algorithm	AZO	2023
Love evolution algorithm	LEA	2024
Human evolutionary optimization algorithm	HEO	2024

4.2 | Physics-Based Methods

Physics-based methods model optimization using physical laws. Post-2020, they focus on binary variants for feature selection. The equilibrium optimizer, published in 2020, is inspired by dynamic mass balance in systems. It initializes particles as subsets, updates equilibrium concentrations, evaluates fitness, and converges. In feature selection, it optimizes multi-objective tasks, as applied in engineering to produce balanced subsets. The Archimedes Optimization Algorithm, from 2021, is inspired by buoyancy principles. It initializes objects, updates via density and volume, evaluates, and iterates. In engineering, it supports feature selection for image fusion. The energy valley optimizer, from 2023, is inspired by energy minimization. It initializes particles, simulates downhill movements, evaluates, and jumps for exploration. In feature selection, it handles global problems for high-dimensional reduction [4], [19].

Table 2 lists these physics-based algorithms, including names, abbreviations, and years.

Algorithm Name	Abbreviation	Publication Year
Equilibrium optimizer	EO	2020
Gradient-based optimizer	GBO	2020
Momentum search algorithm	MSA	2020
Plasma generation optimization	PGO	2020
Transient search optimization	TSO	2020
Black hole mechanics optimization	BHM	2020
Billiards-inspired optimization algorithm	BIO	2020
Archimedes optimization algorithm	AOA	2021
Arithmetic optimization algorithm	AOA	2021
Atomic orbital search	AOS	2021
Chaos game optimization	CGO	2021
Crystal structure algorithm	CryStAl	2021
Lichtenberg algorithm	LA	2021
Material generation algorithm	MGA	2021
String theory-based optimization	STBO	2021
Special relativity search	SRS	2022
Al-Biruni Earth Radius	BER	2023
Energy valley optimizer	EVO	2023
Kepler optimization algorithm	KOA	2023
Young's double-slit experiment optimizer	YDSEO	2023
FATA	FATA	2024
Prism refraction search	PRS	2024
Triangulation topology aggregation optimizer	TTAO	2024
Wave search algorithm	WSA	2024

4.3 | Human-Social Methods

Human-Social methods replicate the cognitive and interactive aspects of humans. Post-2020, they emphasize learning and competition for feature selection. The political optimizer, published in 2020, is inspired by socio-political processes. It initializes parties as subsets, simulates elections and interactions, evaluates fitness, and optimizes. In feature selection, it solves multi-objective problems for engineering.

The mother optimization algorithm, from 2023, is inspired by maternal care. It initializes offspring as subsets, nurtures them for refinement, evaluates them, and matures them. In engineering, it balances dataset selection. The educational competition optimizer, from 2024, is inspired by educational rivalries. It initializes students as subsets, competes for knowledge, evaluates rankings, and refines. Enhances discrete optimization in educational data [8].

Table 3. lists these human-social algorithms with names, abbreviations, and years.

Algorithm Name	Abbreviation	Publication Year
Political optimizer	PO	2020
Gaining and sharing based algorithm	GSK	2020
Human urbanization algorithm	HUA	2020
Interactive autodidactic school	IAS	2020
Nomadic people optimizer	NPO	2020
Buyer-inspired meta-heuristic optimization algorithm	BIMHO	2020
Color harmony algorithm	CHA	2020
FBI-inspired meta-optimization	FBI-MO	2020
Tiki-taka algorithm	TTA	2021
Success history intelligent optimizer	SHIO	2022
Mother optimization algorithm	MOA	2023
Mountaineering team-based optimization	MTBO	2023
Multiple interaction optimizer	MIO	2023
Musical chairs optimization algorithm	MCOA	2023
Running city game optimizer	RCGO	2023
Squid game optimizer	SGO	2023
Skill optimization algorithm	SOA	2023
Special forces algorithm	SFA	2023
Tree optimization algorithm	TOA	2023
Botox optimization algorithm	BOA	2024
Catch fish optimization algorithm	CFOA	2024
Dollmaker optimization algorithm	DOA	2024
Educational competition optimizer	ECO	2024
Election optimizer algorithm	EOA	2024
Football team training algorithm	FTTA	2024
Hiking optimization algorithm	HOA	2024
Human memory optimization algorithm	HMOA	2024
Kids learning optimizer	KLO	2024
Learning cooking algorithm	LCA	2024
Literature research optimizer	LRO	2024
Nizar optimization algorithm	NOA	2024
Object-oriented programming concepts metaheuristic	OOPCM	2024
Partial reinforcement optimizer	PRO	2024
Social small group optimization algorithm	SSGOA	2024

4.4 | Swarm Intelligence Methods

Swarm intelligence methods harness collective dynamics for optimization. Post-2020, they are prevalent for feature selection. The Sparrow Search Algorithm, published in 2020, is inspired by sparrow foraging and anti-predator behaviors. It initializes swarms as subsets, updates via producers/scroungers, evaluates fitness, and optimizes. In feature selection, it handles high-dimensional data for classification.

The marine predators algorithm, from 2020, is inspired by predator-prey interactions in the marine environment. It initializes populations, updates with Lévy/Brownian movements, evaluates, and refines. Applied in feature selection for improved accuracy in classification. The crayfish optimization algorithm, developed in 2023, is inspired by crayfish behavior. It initializes swarms, simulates foraging, evaluates habitat fitness, and optimizes. In feature selection, it supports robust subsets for AI applications [14].

Table 4 lists these swarm intelligence algorithms, including names, abbreviations, and years.

Algorithm Name	Abbreviation	Publication Year
Sparrow search algorithm	SSA	2020
Marine predators algorithm	MPA	2020
Manta ray foraging optimization	MRFO	2020
Mayfly optimization algorithm	MA	2020
Black widow optimization algorithm	BWO	2020
Electric fish optimization	EFO	2020
Sandpiper optimization algorithm	SOA	2020
Slime mould algorithm	SMA	2020
Aquila optimizer	AO	2021
Tuna swarm optimization	TSO	2021
Remora optimization algorithm	ROA	2021
Chameleon swarm algorithm	CSA	2021
African vulture optimization algorithm	AVOA	2021
Dwarf mongoose optimization algorithm	DMOA	2022
Orca predation algorithm	OPA	2022
Pelican optimization algorithm	POA	2022
Prairie dog optimization algorithm	PDOA	2022
Tasmanian devil optimization	TDO	2022
Starling murmuration optimizer	SMO	2022
Sea-horse optimizer	SHO	2022
Honey badger algorithm	HBA	2022
Hunter-prey optimization	HPO	2022
Archerfish hunting optimizer	AHO	2022
Artificial hummingbird algorithm	AHA	2022
Sand cat swarm optimization	SCSO	2022
Poplar optimization algorithm	POA	2022
Komodo mlipir algorithm	KMA	2022
Cheetah optimizer	CO	2022
Dandelion optimizer	DO	2022
Dung beetle optimizer	DBO	2023
Crayfish optimization algorithm	COA	2023
Dark forest algorithm	DFA	2023
Fire Hawk Optimizer	FHO	2023
FOX	FOX	2023
Gazelle optimization algorithm	GOA	2023
Hermit crab optimization algorithm	HCOA	2023
Kookaburra optimization algorithm	KOA	2023
Lyrebird optimization algorithm	LOA	2023
Meerkat optimization algorithm	MOA	2023
Nutcracker optimizer	NO	2023
Osprey optimization algorithm	OOA	2023
Piranha foraging optimization algorithm	PFOA	2023
Prairie dog-based metaheuristic	PDM	2023
Sea-horse optimizer	SHO	2023
Sinh cosh optimizer	SCO	2023
Snow ablation optimizer	SAO	2023
Spider wasp optimizer	SWO	2023
Wolf-bird optimizer	WBO	2023
Addax optimization algorithm	AOA	2024
Arctic puffin optimization	APO	2024
Artificial protozoa optimizer	APO	2024
Black-winged kite algorithm	BWKA	2024
Crested porcupine optimizer	CPO	2024
Dendritic growth optimization	DGO	2024

Table 4. Continued.

Algorithm Name	Abbreviation	Publication Year
Eel and grouper optimizer	EGO	2024
Electric eel foraging optimization	EEFO	2024
Elk herd optimizer	EHO	2024
Flood algorithm	FLA	2024
Geyser-inspired algorithm	GIA	2024
Goose algorithm	GA	2024
Gooseneck barnacle optimization algorithm	GBOA	2024
Greylag goose optimization	GGO	2024
Hippo swarm optimization	HSO	2024
Hippopotamus optimization algorithm	HOA	2024
Horned lizard defense tactics metaheuristic	HLDTM	2024
Leaf in wind optimization	LIWO	2024
Lotus effect optimization algorithm	LEA	2024
Parrot optimizer	PO	2024
Pufferfish optimization algorithm	POA	2024
Puma optimizer	PO	2024
Quokka swarm optimization	QSO	2024
Rhinopithecus swarm optimization	RSO	2024
Secretary bird optimization algorithm	SBOA	2024
Ship rescue optimization	SRO	2024
Snow geese algorithm	SGA	2024
Swarm bipolar algorithm	SBA	2024
Walrus optimizer	WO	2024
Wild gibbon optimization algorithm	WGOA	2024
Yellow ground squirrel algorithm	YGSA	2024

5 | Conclusion

From 2020 to 2025, metaheuristic algorithms have achieved remarkable progress in addressing the NP-hard nature of feature selection. As data complexity and dimensionality have grown, traditional deterministic methods have struggled to cope with exponential search spaces, leading to a surge in adaptive, stochastic, and hybrid metaheuristics. These algorithms have proven highly effective in selecting compact, relevant feature subsets that enhance classification accuracy, reduce redundancy, and improve interpretability across diverse fields from bioinformatics and imaging to cybersecurity and finance.

Over two hundred metaheuristic algorithms were introduced in these five years, reflecting a vibrant era of innovation. Their development can be broadly categorized into four paradigms: evolutionary, physics-based, human–social, and swarm intelligence. Evolutionary algorithms, which account for about 13% of new methods, focus on adaptive evolution and genetic operators that preserve population diversity and generalization. Physics-based algorithms, about 12%, rely on physical modeling principles, such as motion and energy balance, to ensure stable convergence. Human–social algorithms, accounting for roughly 20%, mimic collective human behaviors such as cooperation, competition, and negotiation, offering dynamic adaptability and interpretability. Swarm intelligence algorithms dominate the field, accounting for approximately 55% of new proposals, reflecting their scalability, distributed architecture, and superior performance in high-dimensional feature selection tasks.

This distribution reveals a broader research trend: a shift toward collective, adaptive, and hybrid intelligence. Swarm algorithms are favored for their balance of exploration and exploitation, while human–social methods capture behavioral complexity suited to modern, data-driven systems. Hybrid frameworks that combine multiple paradigms, such as integrating evolutionary diversity with swarm-based refinement or coupling physics-based stability with adaptive learning, have further enhanced performance. Reinforcement learning and chaotic dynamics have also been incorporated, allowing algorithms to self-tune and maintain exploration diversity. The result is a new generation of intelligent optimizers that not only search but also learn, adapt, and self-regulate.

Empirical studies consistently demonstrate that these modern metaheuristics can reduce feature dimensionality by 60–90% while preserving or improving classification accuracy. Their success has established them as a cornerstone of modern machine learning pipelines, enabling faster training, improved generalization, and greater transparency. Moreover, their ability to operate efficiently in high-dimensional and noisy environments positions them as essential tools for real-world applications where scalability and robustness are critical. Nevertheless, challenges remain. Many algorithms depend on random initialization and require careful parameter tuning to achieve optimal results. Ensuring convergence stability, improving computational efficiency, and enhancing reproducibility remain essential research directions. The increasing complexity of hybrid systems also calls for clearer theoretical foundations and standardized evaluation frameworks.

Looking forward, metaheuristic feature selection is moving toward greater intelligence, explainability, and sustainability. Future designs are expected to emphasize interpretable decision processes, online adaptability, energy-efficient computation, and integration with emerging paradigms such as reinforcement learning, transfer learning, and quantum optimization. In summary, the 2020–2025 period represents a transformative era for metaheuristics in feature selection. Evolutionary algorithms contribute adaptive evolution; physics-based methods ensure balance and precision; human–social strategies introduce behavioral intelligence; and swarm intelligence delivers collective scalability. Together, these paradigms have redefined optimization, transforming feature selection from a computational challenge into an intelligent, evolving process that mirrors the principles of learning and cognition.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability

All data are included in the text.

Funding

This research received no specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] Aliyu, D. A., Akhir, E. A. P., Osman, N. A., Salisu, J. A., Saidu, Y., & Yalli, J. S. (2024). Optimization techniques in reinforcement learning for healthcare: a review. *2024 8th international conference on computing, communication, control and automation (ICCUBEA)* (pp. 1–6). IEEE. https://doi.org/10.1109/ICCUBEA61740.2024.10774698

[2] Nagoor, S., & Jinny, S. V. (2023). A dual fuzzy with hybrid deep learning architecture based on CNN with hybrid metaheuristic algorithm for effective segmentation and classification. *International journal of information technology*, *15*(1), 531–543. https://doi.org/10.1007/s41870-022-01106-5

[3] Fakheri, S., Alimoradi, M., & Yamaghani, M. R. (2024). Colour image multilevel thresholding segmentation using trees social relationship algorithm. *Research square*, 1–58. https://doi.org/10.21203/rs.3.rs-4479475/v1

[4] Khaleel, I., Marzoog, W. N., & Al-Kateb, G. (2025). ANILA: adaptive neuro-inspired learning algorithm for efficient machine learning, AI optimization, and healthcare enhancement. *Mesopotamian journal of computer science*, *2025*, 159–171. https://doi.org/10.58496/MJCSC/2025/009

[5] Wang, Y., & Wang, P. (2025). Development and validation of a new diagnostic prediction model for NAFLD based on machine learning algorithms in NHANES 2017-2020.3. *Hormones*, *24*(2), 461–476. https://doi.org/10.1007/s42000-025-00634-6

[6] Olalekan Kehinde, A. (2025). Leveraging machine learning for predictive models in healthcare to enhance patient outcome management. *International research journal of modernization in engineering technology and science*, *7*(1), 1465–1482. https://www.doi.org/10.56726/IRJMETS66198

[7] Yang, Z., Chen, Z., Wang, J., Li, Y., Zhang, H., Xiang, Y., … & Dong, Q. (2025). Multiple machine learning identifies key gene PHLDA1 suppressing NAFLD progression. *Inflammation*, *48*(4), 1912–1928. https://doi.org/10.1007/s10753-024-02164-6

[8] Taji, K., Sohail, A., Shahzad, T., Khan, B. S., Khan, M. A., & Ouahada, K. (2024). An ensemble hybrid framework: a comparative analysis of metaheuristic algorithms for ensemble hybrid CNN features for plants disease classification. *IEEE access*, *12*, 61886–61906. https://doi.org/10.1109/ACCESS.2024.3389648

[9] Alimoradi, M., Azgomi, H., & Asghari, A. (2022). Trees social relations optimization algorithm: a new swarm-based metaheuristic technique to solve continuous and discrete optimization problems. *Mathematics and computers in simulation*, *194*, 629–664. https://doi.org/10.1016/j.matcom.2021.12.010

[10] Alimoradi, M. (2018). *Finding similar batch files with fuzzy clustering*. https://B2n.ir/pd9372

[11] Alimoradi, M., Sadeghi, R., Daliri, A., & Zabihimayvan, M. (2025). Statistic deviation mode balancer (SDMB): A novel sampling algorithm for imbalanced data. *Neurocomputing*, *624*, 129484. https://doi.org/10.1016/j.neucom.2025.129484

[12] Krishnan, P. (2024). Ai-driven optimization in healthcare: machine learning models for predictive diagnostics and personalized treatment strategies. *Well testing journal*, *33*(S2), 10–33. https://welltestingjournal.com/index.php/WT/article/view/Ai_Driven_Optimization_In_Healthcar_Machine_Learning_Models_For_

[13] Alimoradi, M. (2018). *Investigating the composition of apriori algorithm and metaheuristic algorithms (genetic and PSO)*. https://B2n.ir/gj7575

[14] Daliri, A., Asghari, A., Azgomi, H., & Alimoradi, M. (2022). The water optimization algorithm: a novel metaheuristic for solving optimization problems. *Applied intelligence*, *52*(15), 17990–18029. https://doi.org/10.1007/s10489-022-03397-4

[15] Asghari, A., Zeinalabedinmalekmian, M., Azgomi, H., Alimoradi, M., & Ghaziantafrishi, S. (2025). farmer ants optimization algorithm: a novel metaheuristic for solving discrete optimization problems. *Information*, *16*(3), 207. https://doi.org/10.3390/info16030207

[16] Alimoradi, M., Zabihimayvan, M., Daliri, A., Sledzik, R., & Sadeghi, R. (2022). Deep neural classification of darknet traffic. In *Artificial intelligence research and development* (pp. 105–114). IOS press. https://doi.org/10.3233/FAIA220323

[17] Daliri, A., Alimoradi, M., Zabihimayvan, M., & Sadeghi, R. (2024). World hyper-heuristic: a novel reinforcement learning approach for dynamic exploration and exploitation. *Expert systems with applications*, *244*, 122931. https://doi.org/10.1016/j.eswa.2023.122931

[18] Fatima, S. (2024). Healthcare cost optimization: leveraging machine learning to identify inefficiencies in healthcare systems. *International journal of advanced research in engineering technology & science*, *10*(3), 137–147. https://B2n.ir/ry4653

[19] Daliri, A., Branch, K., Sheikha, M., Roudposhti, K. K., Branch, L., Alimoradi, M., & Mohammadzadeh, J. (2023). Optimized categorical boosting for gastric cancer classification using heptagonal reinforcement learning and the water optimization algorithm. *7th international conference on pattern recognition and image analysis (IPRIA)* (pp. 1–6). IEEE. https://B2n.ir/ez9432