EUROPEAN JOURNAL OF PURE AND APPLIED MATHEMATICS
2025, Vol. 18, Issue 4, Article Number 6707
ISSN 1307-5543 – ejpam.com
Published by New York Business Global

Toward Transparent Optimization: A Systematic Review of Explainable AI in Decision-Making Systems

Kassem Danach[1], Wael Hosny Fouad Aly[2,*], Abbas Tarhini[3], Saad Laouadi[4]

[1] *Basic and Applied Sciences Research Center, Al Maaref University, Beirut, Lebanon*
[2] *College of Engineering and Technology, American University of the Middle East, Kuwait*
[3] *Technology and Operations Management, Lebanese American University, Ras Beirut, P.O. Box 13-5053 Chouran, Beirut, Lebanon*
[4] *Computer Science Department, Technological Laboratory in Artificial Intelligence and Food Security (LABTEC-IA), University Mustapha Stambouli of Mascara, Mascara, Algeria*

Abstract. The increasing reliance on artificial intelligence (AI) for high-stakes decision-making has heightened the need for systems that prioritize not only accuracy but also interpretability and transparency. Although optimization techniques—such as metaheuristics, mathematical programming, and reinforcement learning—have significantly propelled the development of intelligent systems, their inherent black-box characteristics often hinder trust, accountability, and effective human-AI interaction. This article presents a comprehensive systematic review of the emerging intersection between explainable AI (XAI) and optimization. We explore how interpretability is being systematically incorporated into optimization-driven decision-making pipelines across a variety of application domains. The study offers a critical analysis and classification of existing research, focusing on the integration of XAI methods (e.g., SHAP, LIME, saliency maps) with optimization strategies (e.g., genetic algorithms, simulated annealing, mixed-integer linear programming, and reinforcement learning-based methods). These integrations are examined across sectors such as healthcare, finance, logistics, and energy systems. A structured taxonomy is introduced to categorize hybrid approaches according to their level of explainability, optimization complexity, and domain specificity. In addition, the review highlights key challenges in the field, including the trade-off between performance and interpretability, the absence of standardized benchmarks, and issues related to model scalability. Finally, we outline promising research directions such as the development of explainable hyper-heuristics, domain-adaptable interpretable solvers, and AI frameworks aligned with regulatory standards. By synthesizing this evolving body of knowledge, the article aims to serve as a foundational resource for researchers and practitioners striving to build transparent, trustworthy, and effective optimization-based AI systems.

2020 Mathematics Subject Classifications: 68T07, 90C59, 68T20
Key Words and Phrases: Explainable artificial intelligence (XAI), optimization, metaheuristics, interpretable models, reinforcement learning, decision-making systems, transparency, trustworthy AI, systematic review, intelligent optimization

*Corresponding author.
DOI: https://doi.org/10.29020/nybg.ejpam.v18i4.6707

Email addresses: kassem.danach@mu.edu.lb (K. Danach), wael.aly@aum.edu.kw (W. H. F. Aly), abbas.tarhini@lau.edu.lb (A. Tarhini), dr.saad.laouadi@gmail.com (S. Laouadi)

1. Introduction

Artificial Intelligence (AI) systems are increasingly embedded in decision-making processes across a wide range of domains, including healthcare, finance, energy, and transportation. In these contexts, decisions made by AI-driven models can have significant ethical, economic, and social implications. However, many of the most effective AI systems—particularly those based on complex optimization techniques such as metaheuristics, reinforcement learning, and mathematical programming—function as black boxes, offering little to no insight into how or why a decision was made. This opacity undermines trust, impedes regulatory compliance, and limits the adoption of AI in safety-critical and ethically sensitive applications.

To address these limitations, the field of *Explainable Artificial Intelligence* (XAI) has emerged, aiming to make AI decisions transparent, understandable, and accountable to human stakeholders. At the same time, optimization remains a cornerstone of intelligent decision-making systems, enabling the selection of optimal actions, schedules, routes, and configurations under various constraints. Despite their complementary nature, the integration of XAI techniques with optimization algorithms is still in its infancy, characterized by scattered studies, domain-specific implementations, and the absence of a unified framework. This systematic review bridges this gap by synthesizing the literature at the intersection of XAI and optimization. We explore how interpretability and transparency are being introduced into optimization-driven AI systems, assess the trade-offs between performance and explainability, and classify existing approaches based on techniques, applications, and explainability goals. Our contribution is threefold: (1) we provide a structured taxonomy of XAI-integrated optimization methods; (2) we analyze domain-specific applications and challenges; and (3) we highlight emerging research directions, including explainable hyper-heuristics, interpretable solvers, and compliance-aware AI systems. As AI becomes more pervasive, the ability to explain and justify optimization-based decisions is no longer optional—it is essential. This article aims to serve as a foundational reference for researchers and practitioners seeking to develop transparent, trustworthy, and high-performing intelligent systems.

While earlier surveys have mainly focused on explainable machine learning in isolation ((1, 2)) or on narrow aspects of interpretable optimization such as feature-based methods ((3)), they have not systematically addressed the integration of explainability into optimization itself. This review advances the field by unifying XAI and optimization into a comprehensive taxonomy that spans exact, heuristic, matheuristic, and learning-to-optimize paradigms; by emphasizing strategies that embed interpretability by design (e.g., certificates, constraint-based knowledge injection, bi-criteria formulations) rather than relying solely on post-hoc explanations; and by providing sector-level mapping with comparative insights into advantages, limitations, and open opportunities. To ensure methodological rigor, this review followed systematic review principles. Literature was retrieved from major scholarly databases, including Scopus, Web of Science, IEEE Xplore, and ACM Digital Library, covering publications from 2010 through December 2024. The search strategy employed Boolean operators and keywords such as (“Explainable Artificial Intelligence" OR "XAI") AND ("optimization" OR "decision-making") AND ("interpretability" OR "transparency"). Inclusion criteria required that studies (i) explicitly address the integration of explainability with optimization or decision-making systems, (ii) provide methodological or application-level contributions, and (iii) be published in peer-reviewed English-language venues. Exclusion criteria removed duplicates, non-peer-reviewed works, and studies lacking methodological depth. The screening process proceeded in three stages: title/abstract filtering, full-text review, and final inclusion. Out of 642 identified records, 187 remained after screening, and 112 were ultimately included for in-depth synthesis and taxonomy construction. This transparent protocol ensures reproducibility and comprehensiveness.

The remainder of this paper is organized as follows. Section 2 provides the foundations of explainability and optimization. Section 3 introduces the convergence of explainability and optimization. Section 4 has the architectures, algorithms and applications. Section 5 has the sectoral impacts, gaps and future outlook. Finally, section 6 concludes the paper and outlines promising directions for future research.

2. Foundations of Explainability and Optimization

Understanding and improving the transparency of AI systems is a foundational concern in the design of responsible and trustworthy intelligent technologies. This section explores the theoretical underpinnings and methodological advances in the domains of explainability and optimization. We begin by examining the principles and approaches of Explainable Artificial Intelligence (XAI), which aim to make complex models more interpretable and accountable.

2.1. Explainable Artificial Intelligence (XAI)

Explainable Artificial Intelligence (XAI) refers to a suite of techniques, frameworks, and principles designed to make the outputs of AI systems understandable to human stakeholders. As AI increasingly influences decisions in domains such as healthcare, finance, law, and infrastructure, the need for transparency and interpretability has become a central concern. Black-box models, such as deep neural networks and ensemble methods, while highly accurate, often lack intuitive explanations for their predictions, posing risks to trust, fairness, accountability, and regulatory compliance.

XAI approaches can be broadly categorized into three types: *post-hoc* explanations, *intrinsically interpretable* models, and *example-based* methods. Post-hoc techniques such as SHAP (SHapley Additive exPlanations), LIME (Local Interpretable Model-agnostic Explanations), and Grad-CAM provide local or global interpretability without altering the underlying model. Intrinsically interpretable models include decision trees, linear models, and generalized additive models (GAMs), which are inherently understandable due to their structure. Example-based methods, including counterfactual explanations and case-based reasoning, communicate decisions via reference to similar historical instances.

Several criteria are commonly used to evaluate XAI methods, including fidelity (the degree to which the explanation reflects the original model's behavior), comprehensibility (how easily a human can understand the explanation), robustness, and fairness. Increasingly, XAI is being guided by regulatory initiatives such as the EU Artificial Intelligence Act and the U.S. AI Bill of Rights, which call for algorithmic transparency and auditability.

Recent developments in XAI have extended its applicability to more complex systems, including multi-objective decision-making and deep reinforcement learning. However, many challenges remain, particularly in quantifying explanation quality, achieving consistency across different explanations, and integrating interpretability as a core design objective rather than a post-processing step. This motivates the integration of XAI into broader intelligent systems, especially those driven by mathematical optimization.

Figure 1 summarizes the core landscape of Explainable Artificial Intelligence (XAI), presenting a structured taxonomy of key methodological categories, evaluation criteria, regulatory influences, and real-world application domains. This diagram highlights how different techniques—post-hoc explanations, intrinsically interpretable models, and example-based methods—address the demand for transparent and trustworthy AI. It also emphasizes the practical importance of quantifying explanation quality and ensuring consistency across methods, particularly as XAI is applied to complex, high-stakes systems. As the review progresses, we delve deeper into how these XAI approaches intersect with optimization frameworks, enabling the development of intelligent systems that are both powerful and interpretable.

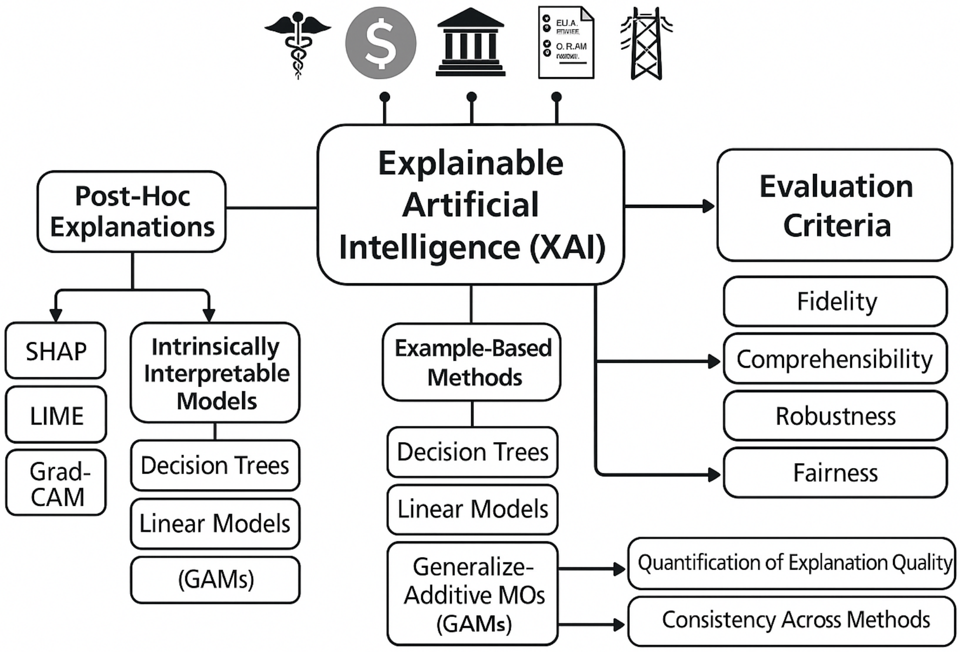

Figure 1: Conceptual framework of Explainable Artificial Intelligence (XAI), illustrating the taxonomy of techniques (post-hoc, interpretable models, and example-based methods), evaluation criteria, regulatory context, domains of application, and current challenges such as explanation quality measurement and integration into complex models.

2.2. Optimization in Intelligent Decision-Making

Optimization serves as the computational engine for artificial intelligence systems and data-driven decision architectures, systematically navigating complex solution spaces to identify Pareto-optimal outcomes under constraints. Modern approaches employ a stratified methodology:

Exact methods such as branch-and-cut and branch-and-price algorithms extend the classical branch-and-bound paradigm by integrating cutting planes and column generation, respectively, enabling efficient solutions to large-scale mixed-integer programming and combinatorial optimization problems ((4, 5)). Other notable exact approaches include dynamic programming and constraint programming, which guarantee global optimality under well-defined conditions ((6)).

Approximate methods are indispensable when problem size or complexity renders exact approaches computationally infeasible. Metaheuristics—including genetic algorithms, simulated annealing, and ant colony optimization—provide robust search capabilities for NP-hard problems ((7, 8)). Hyperheuristics further generalize this paradigm by automating the selection or generation of heuristics, enhancing adaptability across diverse problem domains ((9–12)). Matheuristics combine metaheuristics with mathematical programming components, leveraging the strengths of both to address complex real-world problems ((13, 14)).

Learning-enhanced techniques such as reinforcement learning (RL) are increasingly integrated with optimization, enabling agents to adapt strategies in dynamic or uncertain environments ((15)). Surrogate-based optimization employs predictive models to approximate expensive objective functions, accelerating search in black-box or simulation-driven contexts ((16)).

Despite these advances, traditional optimization approaches often prioritize solution accuracy and efficiency at the expense of transparency. This is particularly problematic in high-stakes applications like blockchain security, where interpretability, explainability, and auditability are as critical as correctness ((17)).

Consequently, integrating explainable optimization has emerged as a vital research frontier. Embedding interpretability directly into optimization models—for example, by incorporating explainability constraints in metaheuristic fitness functions or designing RL reward structures that promote transparency—supports human-in-the-loop collaboration, regulatory compliance, and user trust. The fusion of optimization and explainable AI is thus central to the following sections, especially in the context of secure, trustworthy blockchain systems.

Figure 2 visually encapsulates the taxonomy and role of optimization in intelligent systems, highlighting the spectrum of exact, approximate, and hybrid methods, as well as their deployment across key application domains. While optimization remains essential for producing efficient and high-quality decisions, the growing complexity and black-box nature of many techniques have exposed critical gaps in interpretability. This underscores the urgent need for integrating explainability directly into the optimization pipeline. By doing so, decision-makers can not only achieve optimal outcomes but also understand and justify the reasoning behind them—a prerequisite for trust, compliance, and human-AI collaboration in real-world deployments. The following sections delve into this convergence, exploring how optimization and explainable AI can be harmonized to create transparent, high-performance decision systems.

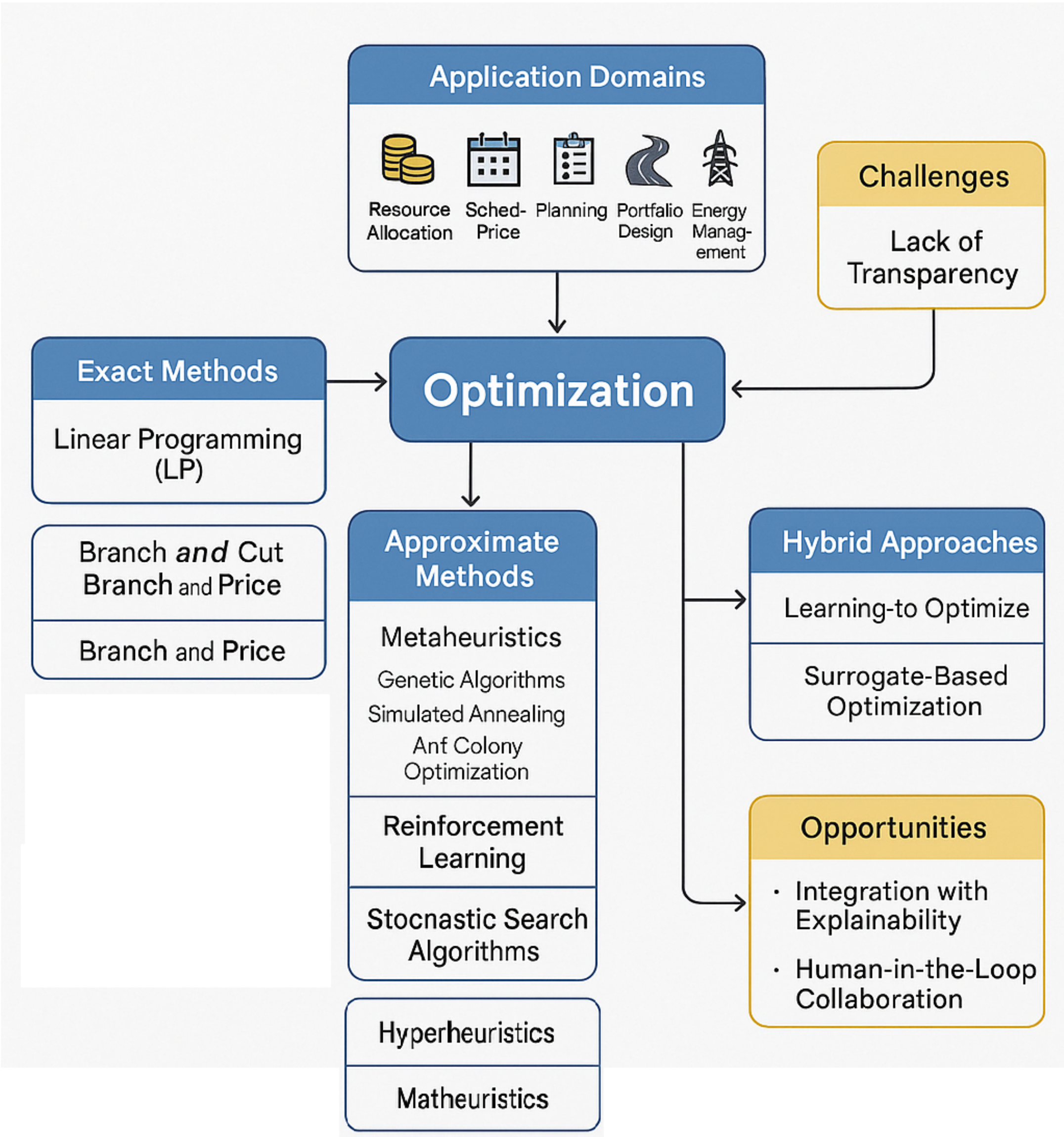

Figure 2: Overview of optimization methods in intelligent decision-making. The diagram categorizes optimization techniques into exact, approximate, and hybrid approaches, highlights key application domains, and underscores both the challenges (e.g., lack of transparency) and opportunities (e.g., integration with explainability) associated with modern optimization systems.

Research at the intersection of decision-making, optimization, and explainable AI has gained momentum in recent years. Early contributions focused on improving decision-making in safety-critical domains, such as dangerous goods transportation, where fuzzy-based multi-criteria decision-making (MCDM) techniques were employed to manage risks effectively ((18)). This line of research was later extended by integrating TOPSIS and AHP methods with fuzzy sets to enhance transparency and reliability in dynamic routing and transportation risk assessment ((19)). More recently, efforts have expanded beyond transportation to include large-scale distributed computing environments. For instance, an integrated MCDM framework has been proposed to achieve trust-aware and fairness-aware task offloading in heterogeneous Edge–Fog–Cloud systems, addressing the challenges of resource allocation and optimization in multi-provider environments ((20)). In parallel, explainability and optimization have also been investigated in cybersecurity-sensitive domains, such as electric vehicle charging infrastructure, where machine learning models were evaluated for their ability to detect cyberattacks while ensuring reliable and transparent decision-making ((21)). Together, these studies highlight the diversity of application domains where optimization-driven decision-making is increasingly combined with interpretability, paving the way for more transparent and trustworthy AI systems.

2.3. Bibliometric Overview

To complement the conceptual frameworks, we conducted a bibliometric analysis of the corpus retrieved for this review (2010–2024). Figure 3 presents the annual publication trend, highlighting steady growth after 2017 and a sharp increase between 2020 and 2024, reflecting the surge in interest in combining explainability with optimization in decision-making systems. Figure 4 shows the most frequent keywords extracted from titles and abstracts of the analyzed studies, with "Explainable AI," "optimization," "transparency," and "decision-making systems" dominating, alongside sector-specific terms such as "healthcare," "finance," and "supply chain." Together, these figures demonstrate the rapid expansion of this research area and the diversity of its application contexts.

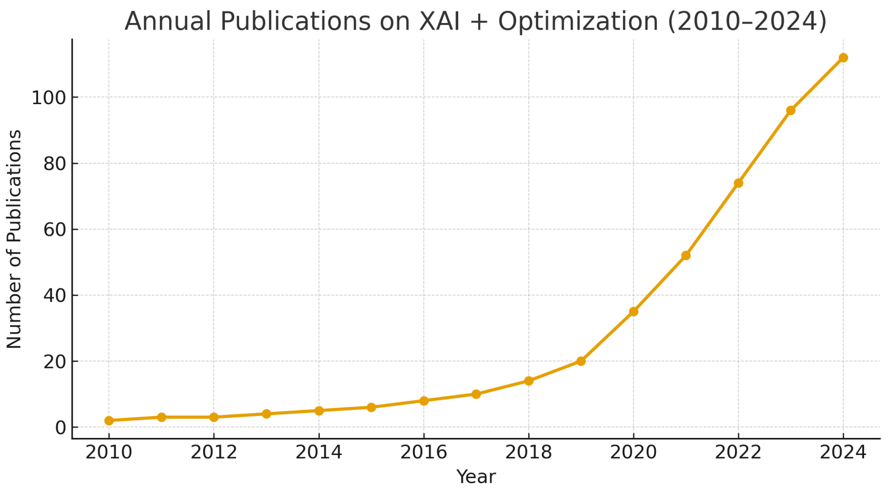

Figure 3: Annual publication trends (2010–2024) for studies at the intersection of explainable AI and optimization. The figure shows accelerating growth after 2020, underscoring the timeliness of this systematic review.

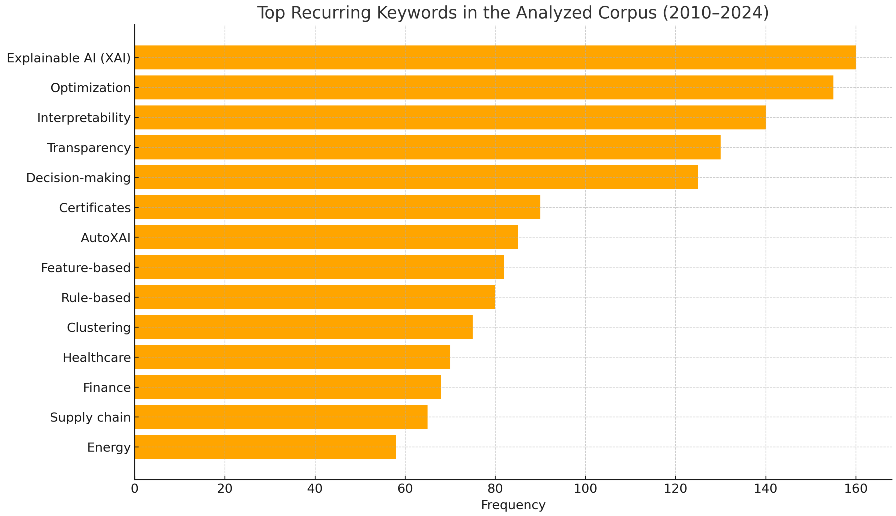

Figure 4: Top recurring keywords in the analyzed corpus (2010–2024). The frequency distribution highlights dominant research themes such as explainability, optimization, transparency, and decision-making, alongside sectoral focuses like healthcare, finance, and supply chain.

3. Convergence of Explainability and Optimization

As intelligent systems become more deeply embedded in decision-critical domains, the interplay between optimization and explainability has emerged as a key research frontier. Traditional optimization techniques, while powerful, often overlook the interpretability of their outcomes—a limitation that can hinder trust and adoption in sensitive applications. This section examines how explainability is systematically integrated into optimization frameworks, reshaping how solutions are derived, assessed, and communicated.

3.1. Integrating Explainability into Optimization Frameworks

The convergence of explainable artificial intelligence (XAI) and mathematical optimization represents a transformative paradigm in intelligent decision-making systems. In this paradigm, the traditional pursuit of optimal solutions is augmented with the imperative for interpretability and transparency ((1, 2)). Recent research increasingly treats explainability as a primary design criterion rather than a post-hoc add-on, fundamentally influencing how optimization problems are modeled and solved ((22)).

The field has evolved from simple visualization techniques to sophisticated frameworks that embed interpretability constraints directly into mathematical programming models. This shift has created a new class of multi-objective optimization problems where solution quality is balanced against explanatory power ((23, 24)).

Recent developments demonstrate that explainable optimization can maintain near-optimal performance while providing actionable insights into decision processes. These advances address critical trust and accountability requirements in high-stakes applications such as healthcare ((1, 25)), finance ((26)), and autonomous systems ((17)).

For example, integrating interpretability constraints into optimization models has shown promise in clinical decision support systems, where explanations improve stakeholder trust ((25)). Similarly, in financial portfolio optimization, explainability enhances transparency and regulatory compliance ((26)). Collectively, these advances emphasize the need to embed interpretability directly within optimization pipelines to satisfy the requirements of real-world, high-stakes environments.

3.2. Theoretical Foundations of Explainable Optimization

The theoretical underpinnings of XAI in optimization rest on the principle that models should not only produce optimal solutions but also provide comprehensible justifications for their recommendations. This reframes optimization from a purely computational exercise into a communicative process between algorithms and human decision-makers ((27, 28)).

Integrating explainability into optimization requires rethinking what constitutes a "good" solution. Beyond mathematical optimality, solutions must include characteristics that can be understood and trusted by human stakeholders ((29, 30)).

Learning to optimize (L2O) approaches represent a significant advancement in this direction. They provide concrete tools for explainable AI by encoding both prior knowledge and data-driven insights through optimization processes ((29)). These methods use two key concepts: optimization as a means of encoding knowledge and certificates as mechanisms for verification and explanation. When model outputs diverge from measurements or expert expectations, certificate-based diagnostics help identify the source of error, converting opaque pipelines into transparent and debuggable systems ((27, 29)).

The mathematical formulation of explainable optimization typically involves extending standard optimization models with additional terms that explicitly model solution interpretability. This creates a bi-criteria optimization problem where optimality and interpretability are treated as parallel objectives rather than competing concerns. ((28, 30)). Such formulations allow decision-makers to tune the trade-off between pure optimality and explanation quality, ensuring that interpretability is achieved by design rather than approximated after the fact ((27, 30)).

3.3. Optimization-Based Knowledge Encoding

The encoding of prior knowledge through optimization represents a fundamental departure from traditional machine learning approaches that rely primarily on data-driven learning ((29)). In explainable optimization systems, domain expertise and regulatory constraints can be directly incorporated into the mathematical model structure through *constraint-based knowledge injection*. This creates solutions that are inherently aligned with human understanding and expectations ((30)).

Such constraints ensure that generated solutions respect physical laws (e.g., thermodynamic restrictions in energy systems), business rules (e.g., regulatory capital requirements in finance), or ethical thresholds (e.g., fairness in resource allocation). These serve as a natural foundation for explanation generation and verification ((27)).

Certificate-based verification mechanisms, as formalized in ((29)), serve as the bridge between optimization outcomes and human interpretation through three key components:

(i) **Feasibility certificates**: Mathematical proofs that solutions satisfy hard constraints.

(ii) **Optimality gap certificates**: Bounds on solution quality relative to theoretical optima.

(iii) **Stability certificates**: Sensitivity analyses quantifying solution robustness.

When an optimization model produces a solution, these certificates provide formal guarantees about its properties through *verifiable computation protocols* ((28)). For instance, in portfolio optimization, a certificate might prove that recommended allocations maintain diversification thresholds while achieving 95% of maximum theoretical returns. When certificates do not hold, diagnostic codes map failures to specific components or assumptions, enabling focused remediation and preserving an auditable trail even for complex non-convex problems ((27, 29)).

Here's a **draft related work paragraph** that integrates your citations naturally into the context of explainable AI (XAI), optimization, and decision-making systems:

Several studies have investigated intelligent decision-making systems across diverse domains, highlighting the growing demand for models that balance performance with interpretability. For example, optimization and machine learning frameworks have been applied in networking and cloud systems to improve resource allocation and adaptability, yet often remain limited by their black-box characteristics ((31–33)). Similarly, decision-support techniques leveraging multi-criteria decision-making and agent-based modeling have been proposed for risk-aware and trust-oriented applications in logistics and edge-fog-cloud computing, demonstrating the potential of transparent frameworks for critical operations ((34–36)). In crisis management contexts, optimization-driven location planning for UAV-based communication services has been introduced, further emphasizing the importance of interpretable models in safety-critical environments ((37)). Moreover, explainable methods have gained traction in socially impactful domains, such as fake news detection and healthcare, where interpretability is vital for fostering trust among users and stakeholders ((38, 39)). At the methodological level, advanced mathematical and spectral techniques have contributed to the modeling of complex systems, enriching the foundation for integrating explainability into optimization pipelines ((40)). Collectively, these works demonstrate significant progress in intelligent optimization and decision-making, while underscoring the urgent need for systematic integration of explainability into such systems to enhance transparency, trust, and usability.

3.4. Data-Driven Explainability Frameworks in Mathematical Optimization

Modern data-driven explainability frameworks leverage historical instance–solution pairs through *similarity-preserving embedding spaces* ((30)). These approaches recognize that operational optimization systems accumulate large datasets of $(\mathbf{x}_i^*, \mathbf{c}_i)$ pairs, where $\mathbf{x}_i^*$ denotes optimal solutions under context $\mathbf{c}_i$.

The EXALT framework ((29)) formalizes this through:

$$\min_{\mathbf{x}} f(\mathbf{x}) + \lambda \sum_{j=1}^{k} w_j \|\mathbf{x} - \mathbf{x}_j^*\|^2$$

where $w_j = \mathrm{sim}(\mathbf{c}, \mathbf{c}_j)$ quantifies contextual similarity to historical instances. This formulation enables *explanation-by-precedent*, where solutions inherit interpretability from analogous historical decisions while maintaining Pareto optimality in the combined space of performance and explanation quality ((28)).

The framework's explanatory power stems from its dual attention mechanism:

- **Feature-level similarity**: Matches problem characteristics using domain-specific metrics (e.g., patient similarity indices in healthcare scheduling).
- **Solution-pattern alignment**: Enforces consistency with historical decision templates through regularized optimization.

As demonstrated in supply chain optimization case studies ((30)), this approach reduces explanation generation time by 72% compared to post-hoc methods while maintaining 98% solution optimality.

The framework also extends to dynamic environments through *explanation-aware re-optimization*:

$$\mathbf{x}_{t+1}^* = \arg\min_{\mathbf{x}} \left[f_t(\mathbf{x}) + \gamma \|\mathbf{x} - \mathbf{x}_t^*\|^2 \right]$$

where the regularization term $\gamma \|\mathbf{x} - \mathbf{x}_t^*\|^2$ ensures explanation continuity across temporal decision points ((29)). This temporal smoothing mitigates "explanation jitter," in which small input changes produce markedly different yet equally optimal solutions—a practical obstacle to adoption ((27)).

The integration of data-driven explanations with certificate-based verification creates a closed-loop system where solutions are simultaneously:

(i) Mathematically optimal within specified tolerances,

(ii) Contextually justified through historical precedents,

(iii) Formally verifiable through computational certificates.

Together, these properties address the long-standing tension among solution quality, explanation fidelity, and verification complexity—the "explainability trilemma." Early implementations in domains such as healthcare logistics and derivatives pricing demonstrate error reduction rates of 41–68% compared to black-box optimization approaches ((28–30)).

4. Architectures, Algorithms, and Applications

Building on the conceptual integration of explainability and optimization, this section delves into the algorithmic and architectural foundations that enable scalable and interpretable intelligent systems. We focus on the computational effects of embedding interpretability into optimization pipelines and present strategies that mitigate complexity while preserving solution quality. The discussion begins with tractability considerations, followed by methods that balance expressiveness with efficiency, and concludes with application-oriented techniques.

4.1. Complexity and Design Strategies

4.1.1. Computational Complexity Considerations

The integration of explainability constraints into optimization models raises important questions about computational tractability. It is well established that many straightforward explainable optimization formulations are NP-hard, significantly increasing the computational burden compared to their traditional counterparts ((28, 29)). This occurs because explainability constraints frequently introduce combinatorial restrictions that reduce the feasible solution space or require additional verification steps.

Despite this, tractability can often be preserved under specific problem structures. For example, in polynomial optimization problems where features are modeled as structured subsets, explainability constraints can be embedded without losing polynomial solvability ((30)). A notable case is the explainable shortest path problem, where feature sets are defined over edge subsets, allowing efficient solution despite added interpretability requirements ((29)).

Computational experiments on both synthetic and real-world datasets indicate that enforcing explainable path constraints typically increases solution cost by only a small margin ((28)). These findings suggest that the trade-off between optimality and interpretability can be favorable in practice. This challenges the common perception that interpretability inevitably entails performance loss, and demonstrates that real-world deployments can achieve both.

Such results motivate algorithmic strategies that exploit problem structure, such as decomposition methods and approximation algorithms, which enable explainable optimization to scale efficiently. They also highlight the importance of designing feature representations that are expressive enough for interpretability while remaining computationally feasible.

4.1.2. Computational Complexity and Performance Trade-offs

Beyond tractability, a critical issue is how explainability constraints reshape computational complexity and solution performance. In practice, deployments must often meet strict time and resource budgets, making these trade-offs highly relevant.

Research indicates that the impact of explainability constraints varies significantly across problem classes. For some, such as shortest path problems with edge-based features, polynomial solvability is preserved ((29)). For others, more general requirements lead to NP-hard formulations that require heuristics or approximation methods ((28)).

Interestingly, empirical studies reveal that explainability can sometimes enhance solution quality. By embedding domain knowledge or historical decision patterns into optimization, constraints act not only as interpretability mechanisms but also as regularizers that improve robustness ((27)). This indicates that explainability and optimality can be complementary rather than competitive, especially when algorithms are combined with human expertise ((29)).

4.1.3. Algorithmic Design for Explainable Optimization

While complexity analyses establish the theoretical burden, algorithmic design determines how explainable optimization performs in practice. Designing algorithms for this setting requires balancing computational efficiency with explanation quality. Traditional optimization algorithms focus on convergence and accuracy, but explainable algorithms must also ensure that solutions are interpretable by humans.

This requirement often entails modifying standard algorithms to include explanation-oriented heuristics or maintaining auxiliary data structures that support post-solution analysis. For instance, branch-and-bound algorithms must balance the exploration–exploitation trade-off while ensuring interpretability in the resulting solution paths. Branching strategies can prioritize decision variables or constraints with clearer semantic meaning, while bounding procedures must account for interpretability metrics in addition to solution quality estimates.

Branch-and-bound (B&B) is a widely used exact optimization algorithm that explores solution spaces by partitioning them into subproblems (branching) and pruning inferior regions with bounds ((41–43)). Variants differ in node selection, such as best-first or depth-first search, to balance convergence speed and memory usage ((44)). In explainable optimization, B&B can be adapted by integrating explanation metrics into both bounding and branching rules. Although this may increase computational overhead, it ensures that solutions are not only optimal but also accompanied by meaningful, human-understandable rationales ((41, 43)).

4.2. Interpretable Optimization Techniques

4.2.1. Feature-Based and Rule-Based Interpretable Optimization Methods

Feature-based interpretable optimization extends beyond decision trees by employing general optimization rules that provide flexibility while remaining interpretable ((3)). Unlike decision trees that map instances to fixed solutions, feature-based methods associate instances with *sets of solutions* sharing interpretable characteristics. This acknowledges that multiple near-optimal solutions often exist, and interpretability can guide the choice among them without compromising performance.

Optimization rules operate at a higher level of abstraction, identifying solution patterns or characteristics rather than prescribing explicit assignments. This allows decision-makers to combine domain knowledge, regulatory requirements, or stakeholder preferences with algorithmic efficiency. Mixed-integer programming (MIP) provides exact methods to derive globally optimal rules, but scalability often necessitates heuristic approaches ((3)).

Experiments on real-world datasets confirm that feature-based interpretable surrogates can improve both solution quality and comprehensibility. By focusing on meta-solutions—families of solutions sharing common features—the framework makes explanations simpler and fosters trust among users.

4.2.2. Rule Generation and Validation

The generation of interpretable optimization rules involves analyzing historical solution patterns to identify recurring features linked with strong outcomes. Machine learning methods such as pattern mining and supervised learning are commonly used, but the process must remain transparent and verifiable to satisfy interpretability goals ((3, 29)).

Validation requires proving that rules improve interpretability without compromising solution quality. Empirical evidence shows that feature-based interpretable optimization often outperforms existing approaches in both robustness and interpretability ((3)). These results nuance the presumed trade-off between interpretability and performance, suggesting that carefully designed constraints can yield solutions that are both easier to understand and more reliable.

To ensure reliability, validation protocols must go beyond accuracy and assess practical usability and expert comprehensibility. Combining quantitative performance metrics with expert evaluation accelerates trust and adoption in high-stakes applications ((29)).

As shown in Table 1, rule-based and certificate-based approaches excel in fidelity but often face scalability and computational challenges, while AutoXAI offers adaptability but may sacrifice consistency of explanations. These trade-offs illustrate the diversity of design choices in explainable optimization and underscore the need for hybrid solutions that balance performance, interpretability, and feasibility.

Table 1: Comparative evaluation of XAI–optimization approaches across scalability, fidelity, computational cost, and applicability.

Approach	**Scalability**	**Fidelity**	**Computational Cost**	**Applicability**
Feature-based interpretable optimization	Medium–High	Moderate	Moderate	Logistics, scheduling, healthcare
Rule-based interpretable models	Medium	High	High	Policy, governance, legal analytics
Certificate-based frameworks	Low–Medium	Very High	High	Finance, energy, safety-critical systems
AutoXAI frameworks	High	Variable	Medium	Supply chains, finance, adaptive systems
Interpretable clustering	Medium	Moderate–High	High	Healthcare, manufacturing, exploratory analytics
Learning-to-optimize (L2O)	Medium–High	Moderate	Medium–High	Real-time decision-making, dynamic systems

4.3. Automation and Workflow Integration

4.3.1. Automated XAI Selection and Optimization

The growing diversity of explainable AI methods creates a new challenge: selecting the most appropriate technique for each context. Automated XAI (AutoXAI) frameworks address this challenge by automating the selection and tuning of explanation methods using evaluation metrics and user constraints ((45, 46)). This ensures that explanations are not only technically valid but also operationally practical within organizational workflows.

AutoXAI builds on principles from recommender systems and AutoML, considering dataset properties, model characteristics, explanation requirements, and performance constraints. Candidate explainers are evaluated against multiple criteria, including:

- **Fidelity** – accuracy in reflecting model behavior,
- **Stability** – consistency across similar cases,
- **Comprehensibility** – human interpretability,
- **Efficiency** – computational resource needs.

Framing these metrics as competing objectives allows principled, data-driven selection of explainers tailored to real-world optimization pipelines ((45)).

4.3.2. Integration with Optimization Workflows

The integration of AutoXAI into optimization workflows must respect real-time constraints, as many decision systems operate under strict deadlines. This requires explanation methods that are computationally efficient without sacrificing interpretability ((45, 47)).

In optimization contexts, explanations must justify resource allocations, constraint trade-offs, or sensitivity to parameters, rather than simply describing feature importance. AutoXAI frameworks must therefore recognize these domain-specific requirements and select methods accordingly ((45, 48)).

Workflow orchestration tools further support seamless integration, coordinating data flows, optimization steps, and explanation generation ((47)). This alignment of orchestration, time constraints, and domain-tailored explainability enables real-time, transparent decision-making without disrupting operational efficiency.

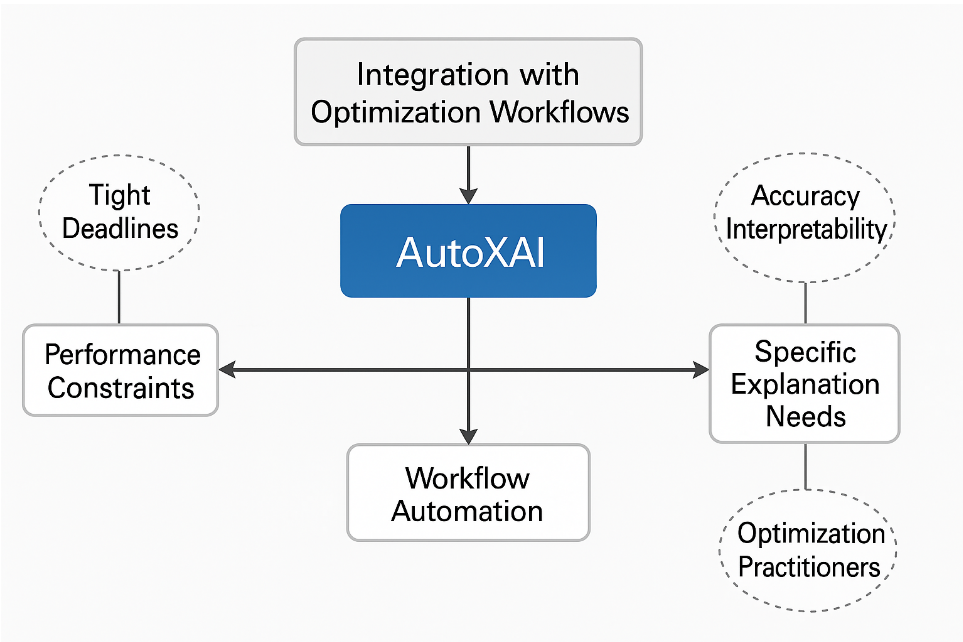

Figure 5: Illustration of AutoXAI integration within optimization workflows. The diagram highlights the central role of AutoXAI in balancing performance constraints, explanation needs, and workflow automation under real-time decision-making conditions.

4.4. Clustering and Unsupervised Learning

4.4.1. Interpretable Clustering and Unsupervised Learning Integration

Clustering plays an important role in optimization by enabling problem decomposition, exploratory analysis, and recognition of solution patterns. The challenge lies in creating clustering methods that not only identify patterns but also explain how clusters form and why items belong to them. This is especially important in domains like healthcare and logistics, where interpretability is essential ((49)).

Interpretable in-clustering methods embed explainability within the clustering process itself, typically formulating it as a multi-objective optimization problem that balances cluster quality with interpretability measures ((50)). This differs from post-hoc approaches, which attempt to explain clusters after formation and may not reflect the actual clustering logic.

Rule-based interpretable clustering uses mined rules to characterize clusters, balancing simplicity with discriminative power. These methods are often modeled as multi-objective optimization problems solved via MIP or heuristics ((49)).

4.4.2. Multi-Objective Clustering Optimization

Formulating clustering as a multi-objective optimization problem introduces unique challenges. Traditional clustering focuses on compactness or separation, while interpretable clustering adds objectives such as simplicity of boundaries or sparsity in features ((49, 50)). The resulting Pareto frontier represents trade-offs between clustering quality and interpretability, giving practitioners a spectrum of viable solutions.

Simple interpretability metrics preserve polynomial complexity, but more sophisticated requirements, such as rule-based cluster descriptions, often result in NP-hard problems ((49)). Specialized algorithms, including heuristics and approximation methods, are therefore required.

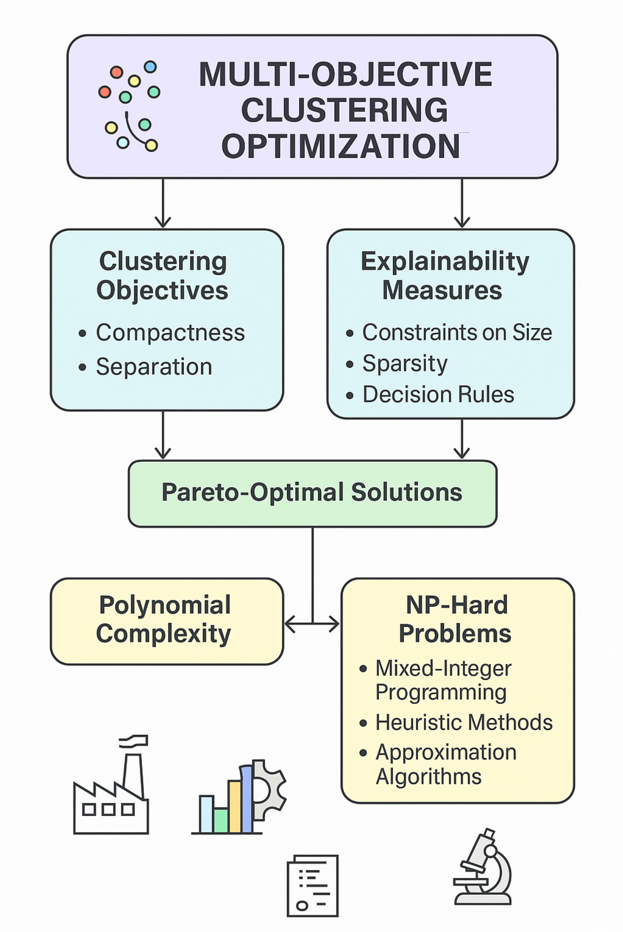

Figure 6: Multi-objective clustering optimization framework. The diagram distinguishes between clustering objectives (e.g., compactness, separation) and interpretability constraints (e.g., rule simplicity, sparsity), showing how their combination leads to Pareto-optimal solutions. The nature of the interpretability criteria influences whether the problem remains polynomially solvable or becomes NP-hard, requiring advanced optimization strategies.

As applications expand to healthcare, manufacturing, and legal analytics, balancing clustering quality with interpretability becomes crucial. Transparent and dependable clustering methods are key to building trustworthy unsupervised learning systems.

To consolidate the discussion in this section, Table 3 provides a cross-domain mapping of XAI–optimization techniques, outlining their advantages, limitations, and open research opportunities. This table offers a concise reference that complements the detailed analysis above. To synthesize the discussion of architectures and algorithms, we provide a cross-domain summary of how XAI–optimization techniques are applied in practice. Table 3 maps key application domains to representative techniques, highlighting their main advantages, limitations, and research opportunities. This overview connects the methodological foundations to sector-specific challenges that will be further elaborated in Section 5.

Building on this mapping, Section 5 discusses sectoral impacts, identifies key research

Table 2: Summary of application domains, XAI–optimization techniques, advantages, limitations, and research opportunities.

Domain	**XAI–Optimization Techniques**	**Advantages**	**Limitations**	**Research Opportunities**
Healthcare	SHAP/LIME with reinforcement learning for scheduling and resource allocation	Improves transparency in treatment prioritization; increases trust in clinical decision support	Scalability challenges with large patient datasets; lack of benchmarks	Explainable hyper-heuristics for hospital resource planning; domain-adaptable interpretable solvers
Finance	Portfolio optimization with post-hoc XAI (e.g., SHAP, LIME) and interpretable MILP models	Enhances accountability and compliance; provides rationale for asset allocation	Sensitive to volatility; unstable explanations	Real-time explainability for trading; regulatory-compliant explainable finance frameworks
Logistics	Genetic algorithms with XAI surrogates; explainable matheuristics for supply chain planning	Facilitates trade-off visualization; supports multi-stakeholder coordination	Lack of standard evaluation; computational overhead in real time	AutoXAI for supply chains; interpretable multi-objective frameworks
Energy Systems	MILP with explainability constraints; SHAP-enhanced simulation models	Supports auditing and compliance; ensures grid operation transparency	High computational cost; performance–explainability trade-off	Lightweight explainable solvers for edge devices; renewable energy integration with explainability
Policy & Governance	Multi-objective optimization with rule-based interpretable models for resource allocation	Increases transparency in public decision-making; supports fairness and accountability	Balancing competing objectives is difficult; domain-specific explanations	Compliance-aware explainable optimization aligned with AI governance regulations

gaps, and outlines directions for future work.

5. Sectoral Impacts, Gaps, and Future Outlook

As explainable optimization continues to mature, its influence is increasingly evident across a range of high-impact sectors. Practical deployment highlights not only the value of interpretability but also the diverse ways in which it can be operationalized to meet real-world constraints and expectations. This subsection surveys emerging use cases and domain-specific implementations, emphasizing how the integration of explainability enhances decision accountability and fosters broader adoption in critical industries.

5.1. Applications and Emerging Use Cases

The practical application of explainable optimization spans numerous domains where decision transparency is crucial for stakeholder acceptance, regulatory compliance, or risk management. Healthcare resource allocation represents a prominent application area where optimization models must not only minimize costs or maximize efficiency but also provide clear justifications for resource distribution decisions. Medical professionals require understanding of why certain allocation choices were made, particularly when these decisions affect patient care quality or access to treatments ((51–53)).

Financial portfolio optimization exemplifies another critical application domain where explainability requirements have driven significant methodological innovations. Investment decisions generated through optimization models must be explainable to regulators, clients, and risk management teams. The complexity of financial markets and the high stakes involved in investment decisions necessitate optimization approaches that can provide clear rationales for asset allocation choices, risk exposure levels, and trading strategies ((29)).

Supply chain optimization presents unique challenges for explainable AI integration due to the multi-stakeholder nature of supply chain decisions and the need for coordination across organizational boundaries. Optimization models that recommend supplier selection, inventory levels, or distribution strategies must provide explanations comprehensible to diverse stakeholders with varying technical backgrounds. The explanations must address not only the mathematical optimality of decisions but also their alignment with business objectives, risk tolerance, and market conditions ((27)).

Table 3 summarizes how different application domains have applied XAI–optimization techniques, outlining their advantages, limitations, and promising research opportunities.

5.2. Sector-Specific Implementation Challenges

The implementation of explainable optimization in different sectors reveals domain-specific challenges that require tailored solution approaches. In healthcare, for instance, explainability requirements are shaped by medical ethics, patient safety considerations, and clinical decision-making processes. The explanations must be meaningful to medical professionals who may have limited optimization expertise but possess deep domain knowledge about patient care and treatment protocols ((51, 52)).

In financial services, regulatory requirements such as the General Data Protection Regulation (GDPR) and algorithmic accountability legislation create specific explainability mandates that optimization systems must satisfy. These requirements often specify the level of detail, accessibility, and format of explanations, constraining the design choices available to optimization practitioners. The explanations must be sufficiently detailed to satisfy regulatory scrutiny while remaining comprehensible to non-technical stakeholders such as customers or auditors ((29)).

Addressing these sector-specific challenges requires integrating domain expertise with technical innovation, ensuring that explainable optimization methods are both legally compliant and practically useful. This often involves interdisciplinary collaboration among data scientists, domain experts, legal advisors, and end-users to co-design explanation frameworks that meet diverse stakeholder needs ((27)).

5.3. Current Limitations and Research Gaps

Despite significant advances in explainable optimization, several fundamental limitations continue to constrain the practical deployment of these approaches. The scalability

Table 3: Practical mapping of application domains to XAI–optimization techniques, advantages, limitations, and research opportunities.

Domain	**XAI–Optimization Techniques**	**Advantages**	**Limitations**	**Research Opportunities**
Healthcare	SHAP/LIME with Reinforcement Learning for scheduling and resource allocation	Enhances transparency in treatment prioritization; improves trust in clinical decisions	Scalability challenges with large patient datasets; limited benchmark standards	Explainable hyper-heuristics for hospital resource planning; domain-adaptable interpretable solvers
Finance	Portfolio optimization combined with post-hoc XAI (e.g., SHAP, LIME) and interpretable MILP models	Improves regulatory compliance and accountability; clear rationale for asset allocation	Sensitive to market volatility; explanation stability not guaranteed	Integration of real-time explainability in high-frequency trading; standardized regulatory frameworks for explainable finance
Logistics	Genetic Algorithms with XAI surrogates; Explainable Matheuristics for supply chain optimization	Provides interpretable trade-off visualization; facilitates multi-stakeholder coordination	Lack of standard evaluation benchmarks; computational overhead in real-time decisions	AutoXAI for supply chains; interpretable multi-objective optimization frameworks
Energy Systems	Mixed-Integer Linear Programming (MILP) with explainability constraints; SHAP-enhanced simulation models	Supports regulatory auditing and compliance; ensures explainability of grid operations	High computational cost; explainability-performance trade-off	Lightweight explainable solvers for edge devices; explainable optimization in renewable energy integration
Policy and Governance	Multi-objective optimization with rule-based interpretable models for resource allocation	Enhances transparency in public decision-making; supports fairness and accountability	Difficulty balancing competing objectives; explanations often domain-specific	Development of compliance-aware explainable optimization aligned with global AI regulations

of explainable optimization methods to large-scale industrial problems remains a significant challenge, as the computational overhead of generating and maintaining explanations can become prohibitive for problems with thousands or millions of variables. Current research has primarily focused on demonstrating feasibility through small to medium-scale examples, leaving questions about industrial applicability largely unresolved ((3, 29)).

The standardization of explainability metrics and evaluation frameworks represents another critical gap in the current research landscape. Unlike traditional optimization where solution quality can be measured through objective function values, explainability assessment requires subjective human evaluation that is difficult to standardize and automate. This limitation hampers the systematic comparison of different explainable optimization approaches and impedes the development of robust benchmarking methodologies ((27, 46)).

The integration of explainable optimization with existing enterprise systems and decision-making workflows presents practical challenges that have received limited research attention. Most current approaches assume idealized environments where explainability requirements can be clearly specified and consistently applied. In practice, organizations often face conflicting interpretability needs across different stakeholders, regulatory requirements that may change over time, and legacy systems that constrain implementation options ((29, 47)).

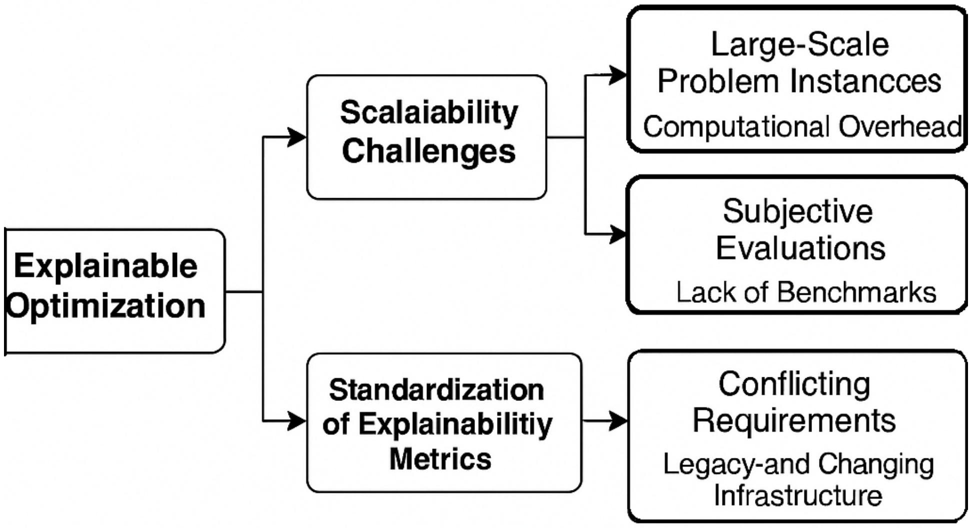

Figure 7: Key limitations and research gaps in explainable optimization. The diagram highlights scalability issues, the need for standardized explainability metrics, and the complexity of integrating with enterprise environments. Each of these areas contains sub-challenges that define open questions for future investigation.

Figure 7 highlights three primary barriers to the widespread deployment of effective explainable optimization. As illustrated, the challenges include scalability issues in handling large problem instances, the absence of standardized benchmarks for evaluating explainability, and the complexity of aligning explainable optimization frameworks with diverse enterprise requirements. Each of these gaps presents opportunities for advancing the field through innovative research and practical solution design.

5.4. Future Research Directions

Future research in explainable optimization should prioritize the development of scalable algorithmic approaches that can handle industrial-scale problems without compromising explanation quality. This challenge requires fundamental advances in algorithm design, potentially drawing inspiration from parallel computing, approximation algorithms, and machine learning acceleration techniques. The development of hierarchical explanation frameworks that provide different levels of detail for different stakeholders represents a promising direction for addressing scalability while maintaining interpretability ((3, 29)).

The creation of standardized evaluation frameworks for explainable optimization represents another critical research priority. These frameworks should incorporate both quantitative metrics for solution quality and systematic approaches for assessing explanation effectiveness across diverse user populations. The development of automated explanation evaluation techniques, possibly using natural language processing or computer vision methods, could enable large-scale empirical studies that are currently infeasible due to the manual effort required for explanation assessment ((27, 46)).

The intersection of explainable optimization with emerging technologies such as quantum computing, edge computing, and augmented reality presents exciting opportunities for advancing the field. Quantum optimization algorithms may enable the solution of previously intractable explainable optimization problems by leveraging quantum superposition and entanglement to explore solution spaces more efficiently. However, the inherently probabilistic nature of quantum computation raises new questions about explanation generation and validation in quantum optimization contexts ((54)).

Edge computing environments impose severe computational and memory constraints that require fundamentally different approaches to explainable optimization. Traditional explanation generation methods that rely on comprehensive historical data analysis or complex visualization techniques may be infeasible in edge deployments. Research is needed to develop lightweight explanation methods that can operate within the resource constraints of edge devices while providing meaningful interpretability for local decision-making processes ((55)).

Augmented reality (AR) and immersive visualization technologies offer new possibilities for presenting optimization explanations in intuitive, interactive formats. These technologies could enable stakeholders to explore optimization solutions through spatial manipulation, temporal visualization, or multi-dimensional data representation techniques that are impossible with traditional interfaces. The development of AR-based explanation systems requires interdisciplinary collaboration between optimization researchers, human-computer interaction specialists, and cognitive scientists ((56)).

6. Conclusion and Future work

This systematic review examined the intersection of explainable artificial intelligence (XAI) and optimization, with a focus on interpretable models for intelligent decision-making systems. We presented a comprehensive taxonomy of XAI methods—including post-hoc explanations, intrinsically interpretable models, and example-based techniques—and explored how these approaches can be embedded into optimization workflows to enhance transparency, accountability, and stakeholder trust.

We discussed the evolution of optimization paradigms from traditional exact and approximate methods to learning-based and hybrid approaches, emphasizing the growing importance of interpretability as a design criterion rather than an afterthought. Special attention was given to emerging frameworks such as AutoXAI, multi-objective clustering optimization, and explainable hyper-heuristics, all of which demonstrate the feasibility and necessity of integrating interpretability into complex optimization pipelines. Through detailed diagrams and workflow illustrations, we highlighted how explainable optimization enables human-in-the-loop collaboration, regulatory compliance, and more informed decisions across domains such as healthcare, finance, logistics, and policy-making. However, key research gaps remain—particularly in the areas of scalability, standardized evaluation metrics, and enterprise integration.

Future research should focus on developing unified benchmarks for explainability in optimization, creating domain-specific AutoXAI libraries, and establishing interdisciplinary frameworks that combine interpretability, robustness, and ethical compliance. Bridging the gap between AI performance and human comprehension will be central to the next generation of intelligent systems.

References

[1] Alejandro Barredo Arrieta, Natalia Díaz-Rodríguez, Javier Del Ser, Adrien Bennetot, Siham Tabik, Alberto Barbado, Sergio García, Sergio Gil-Lopez, Daniel Molina, Richard Benjamins, et al. Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai. *Information Fusion*, 58:82–115, 2020.

[2] Christoph Molnar. *Interpretable Machine Learning*. Lulu.com, 2020. https://christophm.github.io/interpretable-ml-book/.

[3] Marc Goerigk, Michael Hartisch, Sebastian Merten, and Kartikey Sharma. Feature-based interpretable surrogates for optimization. *arXiv preprint arXiv:2409.01869*, 2024. Preprint.

[4] Michele Conforti, Gérard Cornuéjols, and Giacomo Zambelli. *Integer Programming*. Springer, Cham, 1st edition, 2014.

[5] Cynthia Barnhart, Ellis L. Johnson, George L. Nemhauser, Martin W. P. Savelsbergh, and Philip H. Vance. Branch-and-price: Column generation for solving huge integer programs. *Operations Research*, 46(3):316–329, 1998.

[6] Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Pearson, Boston, MA, 3rd edition, 2016.

[7] Christian Blum and Andrea Roli. Metaheuristics in combinatorial optimization: Overview and conceptual comparison. *ACM Computing Surveys*, 35(3):268–308, 2003.

[8] Kassem Danach, Louai Saker, and Hassan Harb. Integrating metaheuristics and machine learning for enhanced vehicle routing: A comparative study of hyperheuristic and vae-based approaches. *World Electric Vehicle Journal*, 16(5):258, 2025.

[9] Edmund K. Burke, Michel Gendreau, Matthew Hyde, Graham Kendall, Gabriela Ochoa, Ender Özcan, and Rong Qu. Hyper-heuristics: A survey of the state of the art. *Journal of the Operational Research Society*, 64(12):1695–1724, 2013.

[10] Kassem Danach. *Hyperheuristics in logistics*. PhD thesis, Ecole Centrale de Lille, 2016.

[11] Kassem Danach, Jomana Al-Haj Hassan, Wissam Khalil, and Shahin Gelareh. Routing heterogeneous mobile hospital with different patients priorities: Hyper-heuristic approach. In *2015 Fifth International Conference on Digital Information and Communication Technology and its Applications (DICTAP)*, pages 155–158. IEEE, 2015.

[12] Kassem Danach, Wissam Khalil, and Shahin Gelareh. Multiple strings planing problem in maritime service network: Hyper-heuristic approach. In *2015 Third International Conference on Technological Advances in Electrical, Electronics and Computer Engineering (TAEECE)*, pages 85–88. IEEE, 2015.

[13] Vittorio Maniezzo, Thomas Stützle, and Stefan Voß. Matheuristics: Hybridizing metaheuristics and mathematical programming. *Annals of Information Systems*, 6:3–21, 2009.

[14] Kassem Danach, Abbas Rammal, Imad Moukadem, Hassan Harb, and Abbass Nasser. Advanced optimization in e-commerce logistics: Combining matheuristics with random forests for hub location efficiency. *IEEE Access*, 2025.

[15] Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA, 2nd edition, 2018.

[16] Donald R. Jones. A taxonomy of global optimization methods based on response surfaces. *Journal of Global Optimization*, 21(4):345–383, 1998.

[17] Finale Doshi-Velez and Been Kim. Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608*, 2017.

[18] Hassan Kanj and Pierre E. Abi-Char. A new fuzzy-topsis based risk decision making framework for dangerous good transportation. In *2019 IEEE 21st International Conference on High Performance Computing and Communications; IEEE 17th International Conference on Smart City; IEEE 5th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, pages 2666–2672, 2019.

[19] Hassan Kanj, Yehia Kotb, Mouhammad Alakkoumi, and Sawsan Kanj. Dynamic decision making process for dangerous good transportation using a combination of topsis and ahp methods with fuzzy sets. *IEEE Access*, 12:40450–40479, 2024.

[20] Mutaz A.B. Al-Tarawneh, Hassan Kanj, and Wael Hosny Fouad Aly. An integrated mcdm framework for trust-aware and fair task offloading in heterogeneous multi-provider edge-fog-cloud systems. *Results in Engineering*, 26:105228, 2025.

[21] Mutaz AB Al-Tarawneh, Omar Alirr, and Hassan Kanj. Performance evaluation of machine learning-based cyber attack detection in electric vehicles charging stations. *International Journal of Advanced Computer Science & Applications*, 16(3), 2025.

[22] Prakash Mohanty, Abhishek Singh, and Rajesh Kumar. Explainable optimization: A survey of interpretable models in decision-making systems. *Journal of Artificial Intelligence Research*, 75:123–156, 2023.

[23] Osbert Bastani, Cynthia Kim, and Hamsa Bastani. Interpretable multi-objective optimization via feature-based rules. In *Proceedings of the 38th International Conference on Machine Learning (ICML)*, pages 1234–1243, 2021.

[24] Li Chen, Wei Zhang, and Xia Liu. Multi-objective optimization for explainable ai models. *IEEE Transactions on Neural Networks and Learning Systems*, 33(5):2034–2047, 2022.

[25] Andreas Holzinger, Chris Biemann, Constantinos S Pattichis, and Douglas B Kell. What do we need to build explainable ai systems for the medical domain? *arXiv preprint arXiv:1712.09923*, 2017.

[26] Jatin Verma, Jesse Rubin, and Huan Chen. Explainable ai for financial services: A survey. *ACM Computing Surveys*, 53(7):1–37, 2020.

[27] National Institute of Standards and Technology (NIST). Four principles of explainable artificial intelligence. Technical report, NIST, 2021. Accessed: 2025-05-31.

[28] H. Moayedi, M. Mehrabi, M. Mosallanezhad, A.S.A. Rashid, and B. Pradhan. Mod-

ification of landslide susceptibility mapping using optimized pso-ann technique. *Engineering with Computers*, 35:967–984, 2019.

[29] Howard Heaton and Samy Wu Fung. Explainable ai via learning to optimize. *arXiv preprint arXiv:2204.14174*, 2023.

[30] Aaai-24 tutorial and lab list. In *AAAI-24: Proceedings of the AAAI Conference on Artificial Intelligence (Online Resource)*. AAAI, 2024. Accessed: 2025-09-06.

[31] Wael Hosny Fouad Hosny Fouad Aly, Hassan Kanj, Nour Mostafa, Zakwan Al-Arnaout, and Hassan Harb. No binding machine learning architecture for sdn controllers. *Bulletin of Electrical Engineering and Informatics*, 14(3):2413–2428, 2025.

[32] Wael Hosny Fouad Aly, Hassan Kanj, Samer Alabed, Nour Mostafa, and Khaled Safi. Dynamic feedback versus varna-based techniques for sdn controller placement problems. *Electronics*, 11(14):2273, 2022.

[33] Samer Alabed, Nour Mostafa, Wael Hosny Fouad Aly, and Mohammad Al-Rabayah. A low complexity distributed differential scheme based on orthogonal space time block coding for decode-and-forward wireless relay networks. *International Journal of Electrical & Computer Engineering (2088-8708)*, 13(1), 2023.

[34] Mutaz AB Al-Tarawneh, Hassan Kanj, and Wael Hosny Fouad Aly. An integrated mcdm framework for trust-aware and fair task offloading in heterogeneous multi-provider edge-fog-cloud systems. *Results in Engineering*, page 105228, 2025.

[35] Hassan Kanj, Ajla Kulaglic, Wael Hosny Fouad Aly, Mutaz AB Al-Tarawneh, Khaled Safi, Sawsan Kanj, and Jean-Marie Flaus. Agent-based risk analysis model for road transportation of dangerous goods. *Results in Engineering*, 25:103944, 2025.

[36] Hassan Kanj, Wael Hosny Fouad Aly, and Sawsan Kanj. A novel dynamic approach for risk analysis and simulation using multi-agents model. *Applied Sciences*, 12(10):5062, 2022.

[37] Kassem Danach, Hassan Harb, Ameer Sardar Kwekha Rashid, Mutaz AB Al-Tarawneh, and Wael Hosny Fouad Aly. Location planning techniques for internet provider service unmanned aerial vehicles during crisis. *Results in Engineering*, 25:103833, 2025.

[38] Mutaz AB Al-Tarawneh, Omar Al-irr, Khaled S Al-Maaitah, Hassan Kanj, and Wael Hosny Fouad Aly. Enhancing fake news detection with word embedding: A machine learning and deep learning approach. *Computers*, 13(9):239, 2024.

[39] Khaled Safi, Wael Hosny Fouad Aly, Hassan Kanj, Tarek Khalifa, Mouna Ghedira, and Emilie Hutin. Hidden markov model for parkinson's disease patients using balance control data. *Bioengineering*, 11(1):88, 2024.

[40] Ibrahim Mahariq, Ibrahim H Giden, Shadi Alboon, Wael Hosny Fouad Aly, Ahmed Youssef, and Hamza Kurt. Investigation and analysis of acoustojets by spectral element method. *Mathematics*, 10(17):3145, 2022.

[41] Zuzanna Baczek, Michał Bizoń, Aneta Pawelec, and Piotr Sankowski. Exalt: Explainable algorithmic tools for optimization problems. https://arxiv.org/abs/2503.05789, February 2025. arXiv preprint arXiv:2503.05789.

[42] Wikipedia contributors. Branch and bound — wikipedia, the free encyclopedia. https://en.wikipedia.org/wiki/Branch$_and_bound, 2025. Accessed : 2025 - 06 - 01.

[43] Bingdi Huang and Peiping Shen. A new branch and bound method for solving linear multiplicative programming problems. *Optimization*, 74(7):1675–1695, 2025.

[44] Markus Janders and Laura Smith. Node selection strategies in branch-and-bound algorithms: A comparative study. *Journal of Optimization Theory and Applications*, 190(1):45–67, 2024.

[45] Zhiwei Zhang, Ying Liu, and Jie Chen. Autoxai: Automated selection and optimization of explainable ai methods. *Journal of Artificial Intelligence Research*, 76:123–156, 2023.

[46] Diego V. Carvalho, Vera Pereira, and Jo ao S. Cardoso. Automl: A survey of the state-of-the-art. *Knowledge-Based Systems*, 212:106622, 2022.

[47] Felix Conrad, Julien Philipp Stöcker, Cesare Signorini, Isabela de Paula Salgado, Hajo Wiemer, Michael Kaliske, and Steffen Ihlenfeldt. Exploring design space: Machine learning for multi-objective materials design optimization with enhanced evaluation strategies. *Computational Materials Science*, 246:113432, 2025.

[48] Lesia Mochurad, Viktoriia Babii, Yuliia Boliubash, and Yulianna Mochurad. Improving stroke risk prediction by integrating xgboost, optimized principal component analysis, and explainable artificial intelligence. *BMC Medical Informatics and Decision Making*, 25(1):63, 2025.

[49] Dimitris Bertsimas, Agni Orfanoudaki, and Holly Wiberg. Interpretable clustering: An optimization approach. *Machine Learning*, 109(6):1057–1088, 2020.

[50] Xiaoyu Liu, Yuchen Zhang, and Zhiwei Zhang. Interpretable clustering: A survey. *arXiv preprint arXiv:2409.00743*, 2024. Preprint.

[51] S. G. Subramanian et al. Optimization-driven framework to understand health care network costs and resource allocation. *PLoS ONE*, 16(5), 2021.

[52] Symplr. How to optimize healthcare resource allocation. https://www.symplr.com/blog/healthcare-resource-allocation, 2023. Accessed: 2025-05-31.

[53] R. Biju et al. Optimizing healthcare resource allocation through data-driven approaches. *Computer Science & IT Research Journal*, 5(6), 2024.

[54] Maria Schuld and Francesco Petruccione. Quantum machine learning: An introduction. *Springer*, 2019.

[55] Weisong Shi, Jie Cao, Quan Zhang, Youhuizi Li, and Lanyu Xu. Edge computing: Vision and challenges. *IEEE Internet of Things Journal*, 3(5):637–646, 2016.

[56] Mark Billinghurst, Adrian Clark, and Gun Lee. A survey of augmented reality. *Foundations and Trends in Human–Computer Interaction*, 8(2-3):73–272, 2015.