Interconnections of Deep Learning, Machine Learning, Reinforcement Learning, ESG Ratings, Artificial Intelligence, Natural Language Processing, and Sustainability: A Survey

Abstract

This survey explores the intricate interconnections between deep learning, machine learning, reinforcement learning, ESG ratings, artificial intelligence, natural language processing, and sustainability. It systematically examines each component's role within the broader context, emphasizing their collective potential to drive innovation and sustainable development. The survey highlights the transformative capabilities of artificial intelligence (AI) and its subsets, including the application of deep learning and machine learning in enhancing precision and scalability across sectors such as healthcare, renewable energy, and education. Reinforcement learning is analyzed for its decision-making prowess in uncertain environments, while ESG ratings are discussed as vital tools for assessing corporate sustainability. Natural language processing's role in facilitating human-computer interaction is underscored, along with its integration with AI for improved language understanding and processing. The survey also addresses the challenges and future directions of these technologies, focusing on ethical considerations, technological limitations, and opportunities for interdisciplinary collaboration. By highlighting the synergies between AI and sustainability, the survey underscores the importance of leveraging these technologies to promote environmentally responsible practices and drive societal advancement. The findings suggest that the integration of AI with ESG and sustainability metrics holds significant promise for optimizing resource management and fostering a more equitable and resilient future.

1 Introduction

1.1 Structure of the Survey

This survey systematically examines the intricate interconnections among deep learning, machine learning, reinforcement learning, ESG ratings, artificial intelligence, natural language processing, and sustainability. It is structured to guide the reader through a comprehensive exploration of each component and its relevance in the broader context. The **Introduction** outlines the significance of these interconnected fields and their increasing importance in contemporary research and industry. The following section, **Background and Definitions**, provides essential definitions and explanations of core concepts, establishing a foundation for deeper analysis.

The survey progresses into **Artificial Intelligence and Its Subsets**, discussing deep learning and machine learning as integral components of AI. The section on **Reinforcement Learning: A Decision-Making Process** highlights its role in decision-making through trial and error. Subsequently, **ESG Ratings and Corporate Sustainability** defines ESG ratings as measures of corporate sustainability, emphasizing their significance. The exploration of **Natural Language Processing for Language Understanding** discusses NLP's capabilities in enabling computers to comprehend human language,

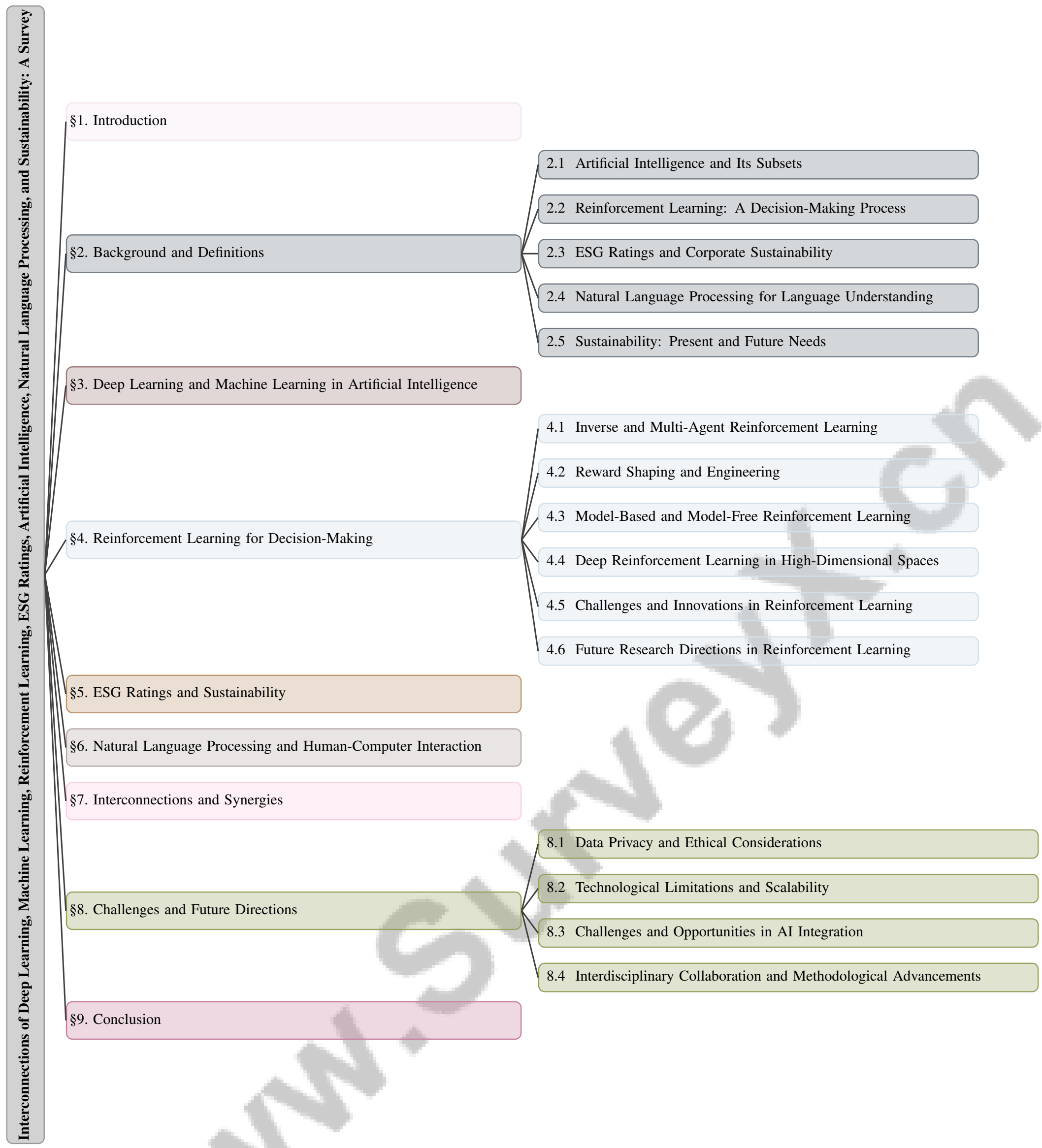


Figure 1: chapter structure

while the section on **Sustainability: Present and Future Needs** concludes foundational discussions by defining sustainability in terms of addressing current and future needs.

The subsequent sections focus on practical applications and challenges of these technologies. **Deep Learning and Machine Learning in Artificial Intelligence** addresses their roles within AI, discussing applications, advancements, and challenges. **Reinforcement Learning for Decision-Making** examines reinforcement learning as a subset of machine learning focused on decision-making, exploring advanced topics and challenges. The section on **ESG Ratings and Sustainability** discusses ESG ratings as tools for assessing corporate sustainability and integrating sustainability into business practices and technological innovations. **Natural Language Processing and Human-Computer Interaction** explores NLP's role in enhancing human-computer interaction, discussing advancements, applications, and challenges.

The survey further analyzes the **Interconnections and Synergies** among these fields, emphasizing the importance of interdisciplinary collaboration and integration strategies. Finally, the **Challenges and Future Directions** section identifies the obstacles in integrating these fields and explores potential future directions, focusing on ethical considerations, technological limitations, and opportunities for interdisciplinary collaboration. The survey concludes by summarizing key findings and underscoring

AI-generated, for reference only.

the potential of these interconnected fields to drive innovation and sustainable development.The following sections are organized as shown in Figure 1.

2 Background and Definitions

2.1 Artificial Intelligence and Its Subsets

Artificial Intelligence (AI) represents a spectrum of computational techniques designed to mimic human cognitive processes, prominently including machine learning (ML) and deep learning (DL). ML is characterized by its ability to create algorithms that learn from data to make predictions, enhancing applications such as financial analysis by improving stock trend forecasts [1]. ML's adaptability is demonstrated through its division into supervised, unsupervised, semi-supervised, and reinforcement learning paradigms, supporting diverse research applications [1].

Deep learning, a specific ML subset, utilizes hierarchical neural networks for feature extraction from complex datasets, eliminating the need for manual feature engineering typical of traditional ML methods [1]. The integration of DL with reinforcement learning is embodied in deep reinforcement learning (DRL), which addresses sequential decision-making challenges [1]. A significant AI challenge is its limited exploratory capability, restricting autonomous knowledge acquisition and adaptation in novel environments [2]. Effective exploration is crucial for data acquisition and knowledge generation, and its absence limits AI's evolution and adaptability [2].

2.2 Reinforcement Learning: A Decision-Making Process

Reinforcement learning (RL) is a pivotal AI component, optimizing decision-making in uncertain environments through a trial-and-error approach. RL agents learn optimal strategies by interacting with their environment and adjusting actions based on rewards or penalties [3]. This iterative method enhances decision-making and adaptability in non-stationary conditions, as seen in HVAC control systems [3].

A primary RL challenge is balancing exploration and exploitation, especially in sparse reward settings. Techniques like intrinsic motivation encourage exploration of new states, addressing the limitations of inadequate external feedback [4]. RL's complexity is compounded by specialized terminology and sophisticated mathematical foundations, presenting challenges for practitioners and researchers [5]. RL's versatility is evident in applications across sectors, such as optimizing microgrid energy conversion under uncertainty [6] and enhancing agricultural management through the CROPS framework [7].

Q-learning is a notable RL algorithm where agents learn optimal actions by evaluating associated rewards or penalties [8]. Despite its effectiveness, challenges like function approximation, sampling error, and nonstationarity can impede performance, necessitating continued research [9]. RL represents a significant AI advancement, offering sophisticated decision-making frameworks in uncertain environments, leveraging elements like value functions and policies, and employing advanced mechanisms such as attention and memory. Its applications span robotics, NLP, healthcare, and finance, with ongoing research to address inherent theoretical and practical challenges [10, 11, 5].

2.3 ESG Ratings and Corporate Sustainability

Environmental, Social, and Governance (ESG) ratings are crucial metrics for evaluating corporate sustainability, providing a comprehensive assessment of performance in these areas. These ratings guide firms toward sustainable practices aligned with societal goals, especially in ML architectures where they evaluate sustainability in energy efficiency, reducing technological infrastructure's environmental footprint [12].

ESG ratings measure corporate sustainability in contexts like distributed energy resources (DERs) integration into energy markets, ensuring economic viability and environmental responsibility [13]. In industries like vehicle fleet optimization, ESG ratings assess efforts to reduce emissions and enhance environmental performance, aligning with sustainability goals [14]. In human-AI collaboration, ESG ratings emphasize integrating ethical considerations into AI systems to promote social equity [15].

AI-generated, for reference only.

ESG ratings are vital for promoting corporate sustainability, integrating environmental, social, and governance considerations into business strategies, particularly in deploying AI and ML technologies. By merging innovation with sustainability objectives, ESG ratings help companies enhance societal well-being and environmental conservation, supporting global biodiversity efforts through initiatives empowering wildlife guardians with advanced technologies for conservation [16, 17, 18, 19, 20].

2.4 Natural Language Processing for Language Understanding

Natural Language Processing (NLP) is foundational in AI, facilitating human-computer interaction by enabling machines to comprehend human language. This capability is critical in Retrieval-augmented Generation (RAG) systems, where NLP bridges preference gaps through effective retrieval mechanisms [21]. NLP's evolution has been accelerated by large language models (LLMs) like LLaMA-2, enhancing applications such as educational feedback systems with nuanced responses [22].

NLP interpretability is assessed through benchmarks that quantify explanation quality, ensuring AI systems provide transparent insights into decision-making [17]. The integration of reinforcement learning with NLP is demonstrated by the Embedding-Aligned Guided Language (EAGLE) agent, refining LLM outputs through reinforcement learning, showcasing the interplay between these techniques [23]. NLP extends to multi-task learning (MTL), collaborating with fields like computer vision and robotics to enhance AI model versatility [24]. Neuro-Symbolic AI (NeSy) in NLP focuses on reasoning, out-of-distribution generalization, and interpretability, critical for learning from small datasets and transferring knowledge across domains [25].

2.5 Sustainability: Present and Future Needs

Sustainability involves meeting present needs without compromising future generations' ability to fulfill their requirements. This principle is crucial in energy forecasting, enhancing resource use efficiency through deep learning techniques applied to energy time series data, supporting sustainable energy management [26].

In transportation, sustainability is addressed by optimizing vehicle speed trajectories for connected and autonomous vehicles (CAVs), minimizing energy consumption and reducing environmental impact [27]. Social challenges like mapping slum locations for effective interventions highlight sustainability's social dimension, advancing Sustainable Development Goals (SDGs) to eradicate poverty and ensure access to basic services [28]. Biodiversity conservation is vital for ecological balance, with local wildlife guardians empowered through economic incentives, linking ecological and economic sustainability [20].

In education, integrating machine learning into civil and environmental engineering curricula equips future professionals with skills to tackle sustainability challenges, emphasizing case studies related to the United Nations SDGs [18]. Sustainability is a multifaceted concept requiring a comprehensive approach that integrates technological, social, and ecological dimensions. This integration is essential for promoting present and future generations' well-being, addressing challenges like biodiversity loss and climate change, and managing the financial and environmental costs of advanced technologies like AI and ML. Innovative tools and methodologies, such as open-source ML toolkits and equitable digital stewardship systems, enable stakeholders to collaboratively develop sustainable solutions [18, 20, 29].

3 Deep Learning and Machine Learning in Artificial Intelligence

The convergence of deep learning and machine learning within artificial intelligence has driven substantial advancements across numerous sectors. This section explores the specific applications and innovations resulting from these technologies, highlighting their transformative impact on fields such as healthcare, renewable energy, and education. As illustrated in Figure 2, the hierarchical structure of these applications reveals a comprehensive overview of data-driven model training and the associated challenges within artificial intelligence. The figure categorizes advancements in human-AI interaction, renewable energy, healthcare, and education, while also emphasizing the critical roles of data, model interpretability, and standardization in the training process. Furthermore, it highlights significant challenges including scalability, ethical considerations, and evaluation standards, underscoring

AI-generated, for reference only.

the necessity for improved interpretability and collaboration to develop effective AI systems. By examining these recent developments, we aim to elucidate the potential of these approaches and their implications in contemporary research and practice.

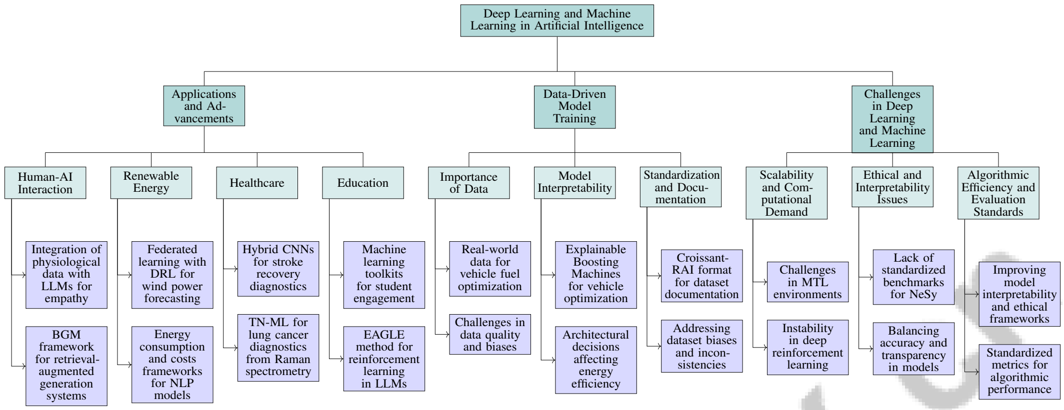


Figure 2: This figure illustrates the hierarchical structure of deep learning and machine learning applications, data-driven model training, and challenges within artificial intelligence. It categorizes the specific advancements in human-AI interaction, renewable energy, healthcare, and education, as well as the importance of data, model interpretability, and standardization in data-driven training. Challenges such as scalability, ethical considerations, and evaluation standards are also highlighted, emphasizing the need for improved interpretability and collaboration for effective AI systems.

3.1 Applications and Advancements

Recent progress in deep learning and machine learning has broadened their applications, underscoring their transformative potential. Advances in AI have improved human-AI interactions by integrating physiological data with large language models (LLMs) to enhance empathy in AI responses [30]. The Bridging the Gap between retrievers and LLMs (BGM) framework exemplifies advancements by seamlessly integrating retrievers with LLMs, enhancing retrieval-augmented generation systems [21].

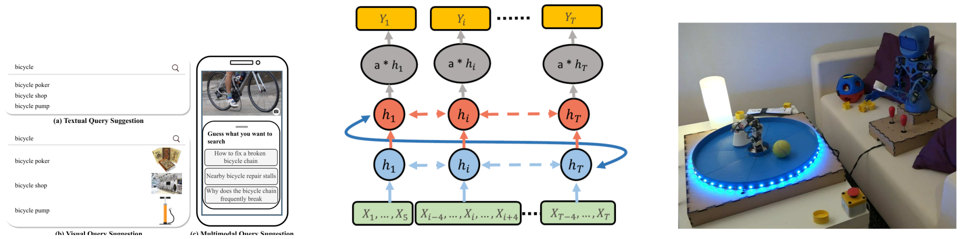
In renewable energy, combining federated learning with deep reinforcement learning (DRL) techniques, such as the deep deterministic policy gradient (DDPG) algorithm, improves wind power forecasting accuracy [31]. This showcases the potential of advanced machine learning in refining prediction capabilities within energy management. Frameworks categorizing the energy consumption and costs of NLP models emphasize the need for efficiency in AI research [29].

In healthcare, hybrid convolutional neural networks in stroke recovery diagnostics integrate MRI-derived regions-of-interest with symbolic clinical data, providing comprehensive insights into patient recovery [32]. Machine learning methods like TN-ML enhance lung cancer diagnostics from Raman spectrometry data, demonstrating the impact of machine learning in medical diagnostics [33].

Explainable AI (XAI) and active learning are crucial for improving human-AI collaboration, highlighting the importance of interpretability [15]. Interpretable models, such as generalized additive models (GAMs), can achieve competitive performance compared to black-box models, challenging the perceived trade-off between accuracy and interpretability [34]. The Explainable Boosting Machine (EBM) advances machine learning applications by predicting fuel consumption and explaining driving behaviors [14].

In education, integrating machine learning toolkits enhances student engagement and understanding of sustainability through hands-on learning [18]. The EAGLE method demonstrates advancements in leveraging reinforcement learning to improve domain-specific task performance of LLMs [23].

The integration of deep learning and machine learning with other AI methodologies continues to drive progress across diverse fields, including natural language processing, where models like BERT and SciBERT enhance language understanding [35]. These advancements pave the way for future innovations addressing complex challenges across multiple sectors, reinforcing the pivotal role of deep learning and machine learning in AI.



(a) Multimodal Query Suggestion for Bicycle Repair[36]

(b) Deep Learning Model with Multiple Layers and Connections[37]

(c) Robotic Interaction with a Ball on a Rotating Platform[38]

Figure 3: Examples of Applications and Advancements

As shown in Figure 3, AI has been transformed by deep learning and machine learning, driving advancements and diverse applications across domains. The multimodal query suggestion system for bicycle repair combines textual and visual data for intuitive search results, enriching user interaction. The depicted deep learning model highlights the intricate architecture of modern AI systems, processing vast data for accurate outputs. Robotic interaction with a ball on a rotating platform exemplifies AI in robotics, where machine learning enables sophisticated environmental interaction, advancing autonomous systems and human-robot interaction [36, 37, 38].

3.2 Data-Driven Model Training

Data-driven model training is crucial in machine learning, where data quality and quantity significantly influence model performance. Training models on real-world data, such as telematics for vehicle fuel optimization, highlights the role of comprehensive datasets in refining model accuracy [14]. Challenges in data-driven training include data quality, availability, and inherent biases.

The deployment of Explainable Boosting Machines (EBM) in vehicle fuel optimization underscores leveraging high-quality data to train interpretable models that provide insights into driving behaviors [14]. Architectural decisions during the inference phase affect energy efficiency and performance, especially in resource-constrained applications [12].

Challenges extend to dataset documentation and standardization. The Croissant-RAI format aims to improve AI system reliability and transparency by standardizing dataset documentation [39]. This addresses dataset biases and inconsistencies that affect model training and evaluation.

Data-driven model training is essential for machine learning progress, requiring collaboration to address challenges related to data quality, documentation, and architectural efficiency. Neural networks, while accurate, can hinder trust due to complexity, especially when training conditions differ from real-world scenarios. Advancements in model interpretability introduce tools clarifying predictions, but scalability issues persist. Effective training methodologies, such as penalizing models for inconsistent explanations, enhance interpretability and generalization across datasets. Teaching key machine learning principles through anti-learning emphasizes validation on unseen data and tailoring approaches to specific challenges [40, 41]. Addressing these issues enables the development of robust models adaptable to diverse real-world scenarios.

3.3 Challenges in Deep Learning and Machine Learning

Deep learning and machine learning encounter challenges in scalability, interpretability, and ethical deployment. One issue is the computational demand for training complex models, especially in multi-task learning (MTL) environments, where large datasets are often required [24]. This is compounded by instability and divergence in deep reinforcement learning, particularly with off-policy learning and bootstrapping [1].

As illustrated in Figure 4, the primary challenges in deep learning and machine learning encompass scalability and efficiency, interpretability and ethics, as well as evaluation and standards. The figure highlights the computational demands placed on models, the difficulties posed by non-stationary environments, the pressing ethical considerations in model deployment, and the critical need for interpretable models alongside standardized metrics for performance evaluation.

Non-stationary environments, such as HVAC systems, challenge traditional heuristic control methods, leading to performance degradation [3]. Adaptive learning algorithms are needed to cope with dynamic changes.

Ethical considerations are critical in deploying machine learning models. The lack of standardized benchmarks for evaluating neuro-symbolic AI (NeSy) complicates performance assessment and trust in AI systems [25]. Interpretability remains a pressing issue, requiring understandable explanations for AI predictions to bridge academic research and practical applications. Evaluating interpretable models against black-box models requires balancing predictive accuracy and transparency. Research shows that generalized additive models (GAMs) can capture complex patterns while remaining interpretable, challenging the notion that only black-box models achieve superior performance. A novel metric for interpretability method quality highlights the need for standardized evaluation criteria, emphasizing understanding model decision-making across applications, including NLP [34, 17].

Addressing challenges in algorithmic efficiency, ethical considerations, and evaluation standards requires improving model interpretability, developing ethical frameworks for AI implications, and creating standardized metrics for algorithmic performance across applications, such as distinguishing human-written from computer-generated texts and enhancing financial analysis with robust datasets and retrieval-augmented techniques [16, 42]. By enhancing data efficiency, improving interpretability, and fostering collaboration, the field can progress toward more effective and trustworthy AI systems.

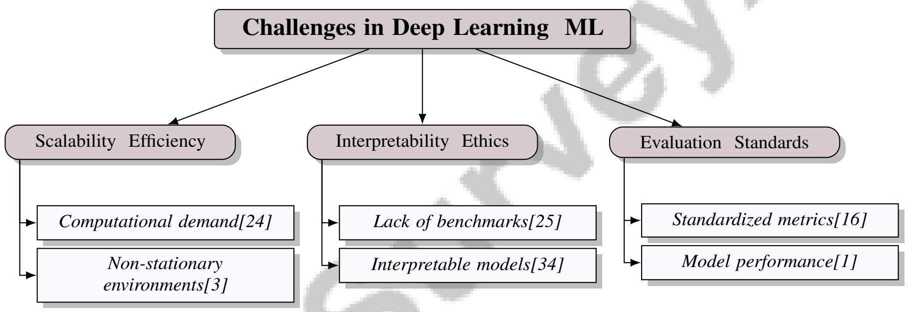


Figure 4: This figure depicts the primary challenges in deep learning and machine learning, including scalability and efficiency, interpretability and ethics, and evaluation and standards. It highlights computational demands, non-stationary environments, ethical considerations, the need for interpretable models, and standardized metrics for performance evaluation.

4 Reinforcement Learning for Decision-Making

Table 1 presents an organized overview of the diverse methods and innovations within reinforcement learning, detailing their categorization, features, and corresponding methodologies to facilitate better understanding and further research in this domain. Additionally, Table 3 provides a comparative analysis of different methods in reinforcement learning, emphasizing their learning paradigms, application domains, and technical innovations, which are crucial for understanding advancements in inverse and multi-agent reinforcement learning. Reinforcement learning (RL) offers robust frameworks for enhancing decision-making capabilities, with Inverse Reinforcement Learning (IRL) and Multi-Agent Reinforcement Learning (MARL) standing out as pivotal methodologies. These approaches address complex decision-making processes, advancing RL's application in intelligent systems across diverse domains.

4.1 Inverse and Multi-Agent Reinforcement Learning

Table 2 presents a comparative analysis of different methods in reinforcement learning, emphasizing their learning paradigms, application domains, and technical innovations, which are crucial for understanding advancements in inverse and multi-agent reinforcement learning. IRL and MARL represent significant advancements in RL, focusing on complex decision-making in intricate environments. IRL deduces reward mechanisms from expert behavior, facilitating the replication of expert strategies in

Category	Feature	Method
Inverse and Multi-Agent Reinforcement Learning	Dynamic Adaptation Strategies	BGM[21]
Reward Shaping and Engineering	Incremental Feedback Mechanisms	ST[43], LCAF[44]
	Human and Data-Driven Adjustments	MA-H-SAC-DF[45], DRLHP[46]
	Contextual and Cognitive Integration	RSM[47], MRL[48], RL-MTVRP[49], ML[50]
Model-Based and Model-Free Reinforcement Learning	Reward and Objective Structuring	IRL[51], PSQD[52]
	Performance Stability	RM-RL[53]
Deep Reinforcement Learning in High-Dimensional Spaces	Temporal Processing	DRNN[54]
	Interpretability Enhancement	AIRL[55]
	Value Function Optimization	RLFA[56]
Challenges and Innovations in Reinforcement Learning	Intervention Optimization	HGRL[57]
	Progressive Learning Strategies	SC[58]
	Efficiency and Scalability	SMR[59], DRL-MOA[60]
Future Research Directions in Reinforcement Learning	Demonstration-Based Approaches	LfND[10]
	Scalability and Efficiency Techniques	AFM[9]

Table 1: This table provides a comprehensive summary of various reinforcement learning methods categorized under key areas such as Inverse and Multi-Agent Reinforcement Learning, Reward Shaping and Engineering, Model-Based and Model-Free Reinforcement Learning, Deep Reinforcement Learning in High-Dimensional Spaces, Challenges and Innovations, and Future Research Directions. Each category lists specific features and methods, along with relevant references, highlighting the latest advancements and methodologies in the field.

Method Name	**Learning Paradigms**	**Application Domains**	**Technical Innovations**
SC[58]	Stepcoder Framework	Code Generation	Curriculum Learning
AFM[9]	Adversarial Feature Matching	Continuous Control Tasks	Adversarial Feature Matching
BGM[21]	-	-	Bridging The Gap
HGRL[57]	Hierarchical Approach	Multi-agent Systems	Graph Neural Networks

Table 2: Overview of innovative methods in reinforcement learning, highlighting their respective learning paradigms, application domains, and technical innovations. This table provides a comparative analysis of various approaches, illustrating their contributions to the fields of inverse and multi-agent reinforcement learning.

autonomous driving and robotics where explicit reward functions are elusive [10, 61]. MARL, on the other hand, deals with multiple interacting agents, balancing cooperation and competition essential for decentralized data fusion systems [15]. Innovations like the Teacher-Student ACL framework enhance learning efficiency in multi-agent contexts [58]. MARL also tackles the credit assignment problem with techniques like Adversarial Feature Matching, improving function approximation and learning performance [9]. Model-based approaches, such as BGM, refine decision-making through adaptive information retrieval [21], while adaptive policies in HGRL show effectiveness in dynamic environments [57]. Cognitive psychology insights into exploratory behavior further enrich RL frameworks [2].

4.2 Reward Shaping and Engineering

Reward shaping and engineering are crucial for enhancing RL by modifying reward signals to improve learning processes. Techniques like the Reward Shaping Method (RSM) integrate contextual information, optimizing agent learning in MARL environments [47]. Step-grained reward shaping provides real-time feedback for complex task execution [43], while metacognitive RL optimizes planning strategies through cognitive process integration [48]. The MA-H-SAC-DF framework addresses imbalanced data environments by tailoring reward structures [45]. The Language-Centric Agent Framework (LCAF) uses language models to facilitate learning in sparse reward settings [44]. Applications in supply chain scenarios demonstrate the efficacy of reward shaping in operational environments [49], with MARLlib enhancing multi-agent settings [50]. Historical approaches iteratively update policies and reward functions through environment interactions and human comparisons [46]. These techniques advance RL by integrating contextual information and employing frameworks like retrieval-augmented generation, enhancing adaptability and robustness [62, 16, 47, 63].

4.3 Model-Based and Model-Free Reinforcement Learning

Model-based and model-free RL are foundational paradigms with distinct advantages and challenges. Model-based RL constructs explicit environmental models for simulation and planning, useful in costly real-world interactions like robotics and urban planning [51]. However, it faces issues like error accumulation and inefficiencies in dynamic conditions [64, 65]. Model-free RL, by contrast, learns optimal policies directly from data, offering robustness against environmental variations [53]. Despite its strengths, model-free RL struggles with data efficiency and generalization, especially with visual inputs, and faces challenges in sparse reward scenarios [62, 66]. Both paradigms contend with state and action space complexity, impacting policy computation [56]. In multi-objective RL, algorithms often lack lexicographic optimization support, leading to suboptimal solutions [52]. Lifelong Learning systems face scalability and resource utilization challenges [67].

4.4 Deep Reinforcement Learning in High-Dimensional Spaces

Deep reinforcement learning (DRL) addresses decision-making in high-dimensional spaces, overcoming traditional RL limitations. Techniques like integrating LSTM cells within RNN frameworks enhance temporal dependency capture in complex data [54]. Adversarial Inverse Reinforcement Learning (AIRL) improves interpretability by summarizing DRL decision-making processes [55]. Sampling techniques combined with function approximation efficiently represent value functions, crucial for scaling DRL in real-world applications like autonomous driving [56].

Figure 5 illustrates the key components of deep reinforcement learning in high-dimensional spaces, highlighting the integration of LSTM cells for temporal dependency capture, the use of adversarial inverse reinforcement learning for improved interpretability, and the application of sampling techniques for efficient value function representation. This visual representation complements the textual discussion by providing a clear overview of how these elements interact within the DRL framework.

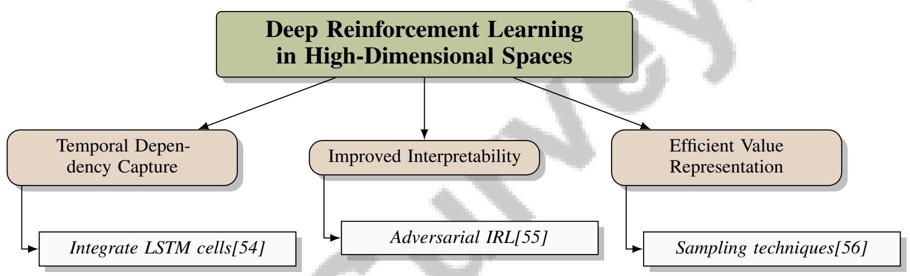


Figure 5: This figure illustrates the key components of deep reinforcement learning in high-dimensional spaces, highlighting the integration of LSTM cells for temporal dependency capture, the use of adversarial inverse reinforcement learning for improved interpretability, and the application of sampling techniques for efficient value function representation.

4.5 Challenges and Innovations in Reinforcement Learning

RL faces challenges in scalability and robustness, particularly in sample inefficiency and generalization [68, 69]. Effective reward function design is critical, especially in sparse reward environments [70, 71]. Integrating intrinsic motivation enhances exploration efficiency [66]. Innovations like Lifelong Learning systems emphasize knowledge retention and adaptation [67], while analog deep learning systems offer high computational efficiency [72]. In MARL, direct punishment mechanisms enhance cooperation [73], and StepCoder improves exploration and optimization accuracy [58]. Off-policy RL algorithms increase sample efficiency [59], with improved performance through addressing function approximation and sampling errors [9]. HGRL simplifies intervention complexities, promoting cooperation [57]. DRL-MOA offers scalability and efficiency advantages over classical MOEAs [60]. Despite challenges in sample inefficiency and reward design, advancements in algorithmic efficiency and lifelong learning highlight RL's potential for scalability and effectiveness, particularly with frameworks like retrieval-augmented generation and large language models [16, 42].

4.6 Future Research Directions in Reinforcement Learning

Future RL research aims to enhance robustness, scalability, and applicability across domains. Improving sample efficiency through meta-learning and integrating human knowledge and auxiliary rewards can optimize exploration methods in complex environments [66, 71]. Unified frameworks integrating solutions across challenges are crucial for real-world testing and practical implementations [74].

In MARL, enhancing scalability and explainability ensures robust decision-making [10]. Further exploration of function approximation methods and hybrid approaches can address off-policy learning limitations [9]. Refining methods for estimating demonstration utility and gathering diverse datasets can improve learning outcomes [10]. These directions aim to address RL challenges and facilitate applications in robotics, natural language processing, finance, and healthcare, enhancing core elements and integrating advanced mechanisms for significant efficacy improvements [63, 11, 71, 1].

Feature	Inverse and Multi-Agent Reinforcement Learning	Reward Shaping and Engineering	Model-Based and Model-Free Reinforcement Learning
Learning Paradigm	Inverse Learning	Reward Modification	Environmental Modeling
Application Domain	Autonomous Systems	Supply Chain	Urban Planning
Technical Innovation	Teacher-Student Acl	Language-Centric Agent	Lifelong Learning Systems

Table 3: This table provides a comparative analysis of three prominent reinforcement learning methodologies: Inverse and Multi-Agent Reinforcement Learning, Reward Shaping and Engineering, and Model-Based and Model-Free Reinforcement Learning. It highlights their distinct learning paradigms, application domains, and technical innovations, offering insights into their unique contributions to the field.

5 ESG Ratings and Sustainability

The integration of environmental, social, and governance (ESG) factors into business practices is increasingly essential, as it aligns corporate strategies with sustainability goals. This section examines the role of sustainability in business practices and highlights technological innovations that drive sustainable development within organizations.

5.1 Sustainability in Business Practices

Incorporating sustainability into business practices is crucial for long-term success and societal welfare. Technologies like machine learning (ML) and artificial intelligence (AI) are pivotal in creating sustainable business models by optimizing decision-making and resource use. Enhanced predictive models, for instance, improve efficiency in automated trading and corporate decision-making [19]. In energy management, reinforcement learning in hybrid electric vehicles (HEVs) boosts fuel efficiency and reduces emissions, aligning with sustainability objectives [75]. Smart building technologies, such as adaptive HVAC systems, further showcase the role of AI in promoting energy efficiency [3]. Explainable Boosting Machines (EBM) optimize driving behavior, leading to cost and emissions reductions [14]. Comparative benchmarks of CNN acceleration methods in edge computing emphasize energy efficiency, supporting sustainable AI processing [76].

Human-centered mechanism design aligns economic policies with societal preferences, fostering equitable economic strategies that support sustainability [77]. This approach underscores the potential of aligning business strategies with sustainable development goals, balancing economic growth with environmental responsibility.

Addressing challenges like biodiversity loss and resource depletion requires leveraging technology and local conservation efforts [29, 16, 18, 19, 20]. By integrating technological innovations and aligning business strategies with sustainability objectives, companies can enhance societal well-being and environmental conservation.

5.2 Technological Innovations and Sustainability

Technological innovations are crucial for sustainability, offering solutions that enhance efficiency and reduce environmental impacts across various sectors. The RL4Sugg framework, for example, uses reinforcement learning to optimize user engagement in search engines, promoting sustainability through resource-efficient operations [36]. In recycling, CNN-based systems improve material sorting, enhancing recycling efficiency and reducing environmental impact [78].

In the automotive sector, the NN-Rollout method provides computational savings and memory efficiency for real-time eco-driving applications, contributing to emissions reduction and energy efficiency [27]. The integration of IoT and big data technologies further improves service flexibility and energy management [79]. Platforms like TensorFlow support sustainable technology development by enabling efficient machine learning model creation [80]. Edge AI applications address global challenges, demonstrating technology's role in promoting sustainability by meeting essential human needs [81].

Challenges persist, such as the high costs of specialized hardware and the environmental impact of non-renewable energy consumption [29]. Addressing these requires developing energy-efficient computational methods and leveraging renewable energy sources.

Technological innovations advance sustainability by improving efficiency, reducing resource consumption, and addressing societal challenges. The ongoing development of ML and AI is crucial for creating sustainable practices aligned with the United Nations Sustainable Development Goals (SDGs). These technologies automate complex tasks in environmental engineering, providing valuable data for informed decision-making. Approaches like digital stewardship for biodiversity conservation empower local wildlife guardians, ensuring ecosystem protection and equitable resource distribution [18, 20, 29, 28].

5.3 Challenges in ESG and Sustainability Metrics

Benchmark	Size	Domain	Task Format	Metric
CoinRun[82]	2,000,000	Reinforcement Learning	Generalization	Generalization Metric
ChatGPT-DT[83]	10,000	Natural Language Processing	Text Classification	Accuracy, F1-score
SG-RH-AC[84]	10,000	Autonomous Driving	Object Detection	SG, RH
GBT[85]	3,510	Linguistics	Bias Detection	Krippendorff's alpha
ZSL-LLM[86]	4,000	Sentiment Analysis	Text Classification	F1 Score, Accuracy
EDA[87]	1,000	Entity Deduction	Entity Deduction Game	Exact Match, Success Rate
TF[80]	1,000,000	Computer Vision	Image Classification	Accuracy, F1-score
ML-BENCH[88]	9,641	Machine Learning	Task Execution	Pass@5

Table 4: This table presents a comprehensive overview of various benchmarks utilized in different domains, highlighting their size, domain of application, task format, and evaluation metrics. The benchmarks span across diverse fields such as reinforcement learning, natural language processing, autonomous driving, linguistics, sentiment analysis, entity deduction, computer vision, and machine learning. This collection serves as a foundational reference for understanding the scope and evaluation criteria of benchmarks relevant to AI and machine learning advancements.

ESG and sustainability metrics face challenges that affect their efficacy and reliability. A major issue is the substantial computational resources needed for large language models (LLMs), which limits their practical application and scalability [22]. This highlights broader challenges in deploying AI for ESG metrics, where resource-intensiveness hinders adoption.

Bridging the preference gap between retrievers and LLMs complicates effective ESG and sustainability metric development [21]. This reflects a broader issue of aligning AI models with ESG assessment requirements, where integrating diverse data sources and preferences is challenging.

Multi-task learning integration presents challenges like effective knowledge transfer and model optimization for diverse requirements [24]. These are particularly relevant for ESG metrics, where synthesizing information across domains is crucial for accurate assessments.

Addressing these challenges involves enhancing algorithmic efficiency, improving AI system transparency, and establishing robust benchmarks for ESG metrics. Table 4 provides a detailed overview of representative benchmarks that are critical for addressing the challenges in ESG and sustainability metrics, particularly in the context of enhancing algorithmic efficiency and establishing robust evaluation frameworks. Leveraging machine learning advancements and interpretability methods, alongside integrating diverse data sources, ensures models are accurate and provide clear reasoning for predictions. Developing comprehensive frameworks that incorporate real-time data and promote AI decision-making understanding aligns technological advancements with sustainability goals [16, 18, 17, 42]. Overcoming these obstacles can improve ESG metric reliability and applicability, fostering more effective and sustainable business practices.

5.4 Future Directions in ESG and Sustainability

The future of ESG and sustainability initiatives will be shaped by integrating advanced technologies and comprehensive regulatory frameworks. Emphasizing robust regulatory frameworks is key to addressing social issues and ensuring technological developments align with environmental stewardship [89]. Figure 6 illustrates these future directions in ESG and sustainability, focusing on the integration of advanced technologies, the development of regulatory frameworks, and the refinement of data methodologies. Each category highlighted in the figure underscores critical areas such as AI applications, ethical considerations, and context-specific data models.

Future research should focus on context-specific models and improved data collection methods, including synthetic datasets for urban challenges like slum mapping [28]. Refining these models allows policymakers to implement effective interventions aligned with sustainable development goals.

Expanding life cycle assessments is critical for evaluating environmental impacts associated with technologies. Future research should broaden these assessments to include more computing technologies, fostering a sustainable approach to AI processing [76]. This will enable informed decisions about technology deployment, ensuring environmental considerations are integral to innovation.

Future ESG and sustainability directions hinge on integrating advanced technologies with ethical and regulatory considerations. By prioritizing context-specific solutions, refining data methodologies, and expanding life cycle assessments, these initiatives can enhance sustainable development efforts, fostering an equitable and environmentally conscious society. Leveraging ML and AI for efficient data processing and analysis addresses issues like biodiversity conservation and sustainable resource management. Integrating these approaches into education and community engagement empowers stakeholders to participate in conservation efforts, ensuring sustainable practice benefits are shared equitably [18, 20, 29, 28].

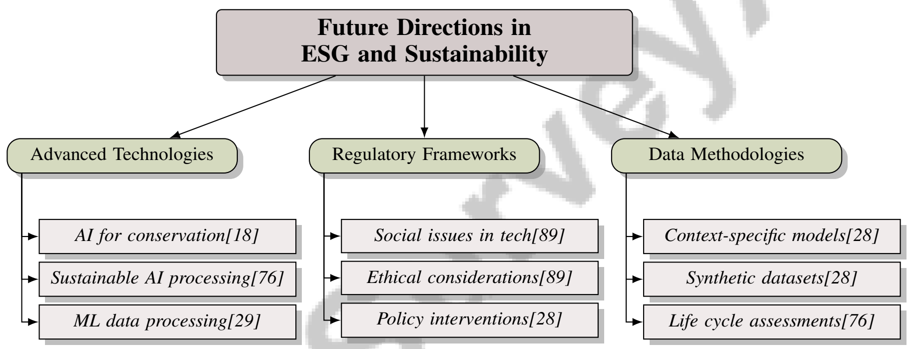


Figure 6: This figure illustrates the future directions in ESG and sustainability, focusing on the integration of advanced technologies, the development of regulatory frameworks, and the refinement of data methodologies. Each category highlights key areas such as AI applications, ethical considerations, and context-specific data models.

6 Natural Language Processing and Human-Computer Interaction

The interplay between Natural Language Processing (NLP) and Human-Computer Interaction (HCI) is pivotal, driven by NLP advancements that transform user engagement with digital systems. This section delves into NLP progress, its applications, and implications for enhancing user interactions, culminating in the discussion of advancements that facilitate sophisticated communication between humans and machines.

6.1 Advancements in NLP Technologies

Recent NLP advancements have significantly improved language understanding and generation tasks. Deep Recurrent Neural Networks (RNNs), particularly with Connectionist Temporal Classification (CTC) and RNN transducers, have enhanced speech recognition [54]. Benchmarks for evaluating multi-turn planning in large language models (LLMs) like GPT-4, GPT-3.5, Claude-1/2, and Vicuna highlight their planning and decision-making capabilities, essential for developing complex conversational agents [87]. AUTONLU models have set new performance benchmarks across various datasets, reflecting NLP's continuous evolution [35].

The Interpretation Quality Score (IQS) introduces a metric for assessing NLP models' interpretability, fostering trust in AI systems [17]. These advancements underscore NLP's rapid progression and

AI-generated, for reference only.

transformative potential for HCI. By integrating deep learning and symbolic reasoning, NLP enhances AI systems' accuracy, efficiency, and interpretability. Neuro-symbolic AI addresses limitations in abstract reasoning and out-of-distribution generalization, while frameworks like AutoNLU streamline natural language understanding model development [25, 35, 54, 17].

6.2 NLP Applications in Human-Computer Interaction

NLP enhances HCI by enabling intuitive human-machine communication. Intelligent virtual assistants and chatbots utilize advanced NLP techniques for seamless interactions across platforms [35]. In education, NLP transforms grading and feedback systems by employing LLMs to assess submissions and provide personalized feedback, enhancing learning experiences [22]. In healthcare, NLP facilitates patient interactions with natural language interfaces, improving service accessibility and efficiency [33].

NLP technologies also promote accessibility, developing assistive tools for individuals with disabilities, such as deep learning-powered speech recognition systems for voice-controlled interfaces and real-time transcription [54]. These advancements illustrate NLP's role in fostering user-friendly interfaces and enhancing communication capabilities [25, 35, 42, 54].

6.3 Challenges in NLP: Bias and Fairness

Bias and fairness are critical challenges in NLP, affecting AI technologies' reliability and ethical implications. Biases in LLMs can perpetuate stereotypes, necessitating vigilant oversight during model training [90]. The instability of fairness metrics, especially with small sample sizes, complicates model fairness assessments [91]. Discrepancies in interpretability methods underscore the need for robust frameworks to reconcile varying interpretations, enhancing transparency [92].

As illustrated in Figure 7, the primary challenges in NLP related to bias and fairness are depicted, emphasizing the multifaceted nature of these issues. This figure highlights the biases present in large language models, the problems associated with fairness metrics, and the critical importance of interpretability. Each category delineates specific aspects, such as the reinforcement of stereotypes, the instability of fairness metrics, and the necessity for personalized explanations.

Addressing these challenges requires personalized explanations that consider individual user backgrounds, mitigating biases and enhancing fairness [93]. Strategies involve improved evaluation metrics, personalized interpretability, and ethical AI practices to develop equitable and trustworthy systems [91, 85, 42, 17, 35].

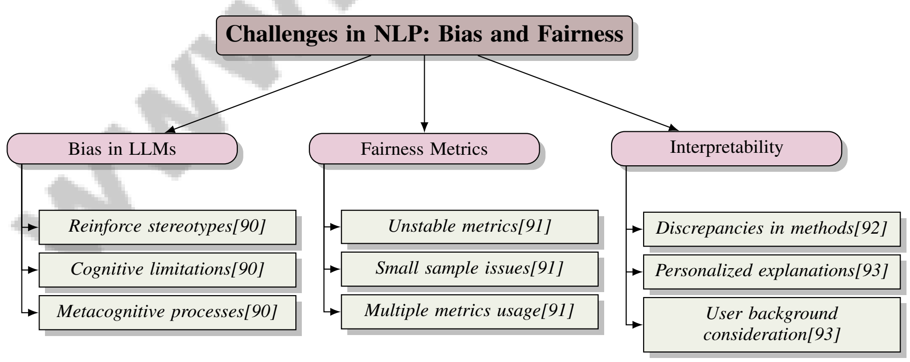


Figure 7: This figure illustrates the primary challenges in NLP related to bias and fairness, highlighting biases in large language models, issues with fairness metrics, and the importance of interpretability. Each category includes specific aspects such as reinforcing stereotypes, unstable metrics, and the necessity for personalized explanations.

6.4 Interpretability and Reasoning in NLP

Interpretability and reasoning are essential in NLP model development, enhancing transparency and user trust. Generating global summaries of model behavior, as seen in interpretable deep reinforcement learning, emphasizes creating comprehensive and user-friendly explanations [55].

Interpretability aligns AI systems with ethical standards and user expectations, enabling informed decisions based on AI insights.

Reasoning capabilities enhance NLP models' utility by enabling logical inferences and contextually relevant decisions, vital for applications like conversational agents and automated decision-making systems. Integrating robust reasoning frameworks with NLP models improves their ability to interpret and respond to ambiguous inputs. Neuro-symbolic AI combines deep learning with symbolic reasoning, addressing traditional deep learning systems' limitations in abstract reasoning, thus facilitating improved understanding and interpretability in applications such as financial analysis [16, 25, 94, 17].

6.5 NLP and Sustainability

The intersection of NLP and sustainability addresses aligning technological advancements with environmental and ethical considerations. As NLP models, particularly LLMs, become more sophisticated, their computational demands raise concerns about environmental impact. Developing energy-efficient algorithms and hardware is essential for sustainable AI practices [29]. Strategies for enhancing sustainability in NLP include transparent reporting of training times and computational costs, promoting awareness and efficient methodologies. Equitable access to computational resources ensures inclusive NLP advancements [29].

Innovative approaches explore eco-friendly model architectures that prioritize computational efficiency without sacrificing performance. Incorporating sustainability principles into NLP design mitigates environmental challenges posed by large-scale model training, promoting a responsible future for the field. This aligns with findings on the significant energy consumption of modern deep learning techniques, advocating for eco-friendly practices in AI development [18, 17, 29, 35]. Prioritizing energy efficiency, equitable resource distribution, and transparent reporting will enable the NLP community to contribute to a sustainable and ethically responsible future.

7 Interconnections and Synergies

Examining the interplay between technological advancements and sustainable practices highlights the synergies that arise when artificial intelligence (AI) intersects with sustainability initiatives. This exploration underscores AI's transformative potential and the necessity of collaborative efforts to leverage these technologies for sustainable development. The following subsection elucidates the synergies between AI and sustainability, showcasing how these fields collectively address pressing global challenges.

7.1 Synergies between AI and Sustainability

The synergy between AI and sustainability is evident as technological advancements address critical environmental and social challenges. Applications of deep reinforcement learning demonstrate AI's capability to manage complex environments, automate feature extraction, and enhance resource management, thereby reducing environmental impact [1]. The EAGLE framework exemplifies how aligning large language model (LLM) generation with latent embedding objectives optimizes content generation and user satisfaction, minimizing resource wastage [23].

AI's role in sustainability extends to interdisciplinary applications, such as deep learning for urban planning and policy-making, providing data and insights essential for informed decision-making on social and environmental issues. This intersection necessitates interdisciplinary collaboration to harness AI's capabilities for sustainable development. Ensuring transparency and explainability in AI systems is crucial for building stakeholder trust and promoting sustainable practices, aligning initiatives with sustainability goals to foster environmental stewardship and social equity. Addressing algorithmic biases and inconsistencies ensures technological advancements contribute positively to biodiversity conservation and responsible resource management [39, 18, 20, 95, 96].

Integrating AI advancements with sustainability initiatives offers a transformative approach to achieving global sustainability goals. Engineers can utilize machine learning tools for efficient environmental monitoring and resource management while addressing financial and ecological costs associated with AI deployment through solutions like edge computing and energy-efficient hardware

[76, 18, 29]. Leveraging AI's potential to optimize resource utilization and drive innovation presents an opportunity to foster a more equitable and resilient future.

7.2 Interdisciplinary Collaboration and Integration

Interdisciplinary collaboration is vital for maximizing AI and machine learning (ML) potential to tackle complex challenges across domains. Integrating AI, sustainability, and human-computer interaction requires leveraging each discipline's strengths. Collaborations among AI, finance, and data analytics researchers can yield innovative solutions using advanced methodologies like deep learning and large language models, adaptable to real-world scenarios and enhancing interpretability in complex decision-making processes [16, 19, 81, 29].

A crucial area for collaboration is developing AI systems aligned with sustainability goals. Integrating AI with environmental sciences enhances model accuracy and effectiveness in climate change mitigation and resource management, optimizing energy consumption and minimizing environmental impact [29]. The intersection of AI and social sciences provides insights into AI technologies' ethical and societal implications, addressing concerns related to bias, fairness, and transparency. Incorporating ethics, sociology, and psychology perspectives ensures AI technologies are socially responsible and aligned with human values [92].

In human-computer interaction, collaboration between AI researchers and user experience design experts creates intuitive and accessible interfaces, enhancing AI systems' usability across various contexts and abilities. Prioritizing user-centered design allows interdisciplinary teams to develop functional and user-friendly AI applications [54].

As illustrated in Figure 8, interdisciplinary collaboration in AI encompasses its integration with sustainability, social sciences, and human-computer interaction, addressing complex challenges while enhancing usability and ethical standards. This collaboration is essential for integrating AI and ML with diverse fields, facilitating standardized practices, enhancing interpretability methods, and supporting innovative applications in real-time environments like IoT and edge computing. It addresses data quality, interoperability, and responsible AI requirements, leading to efficient and impactful technologies across domains [81, 97, 39, 17, 24]. Combining expertise from various disciplines allows researchers to devise comprehensive solutions addressing modern society's multifaceted challenges, advancing technological innovation while aligning AI systems with societal goals and ethical standards.

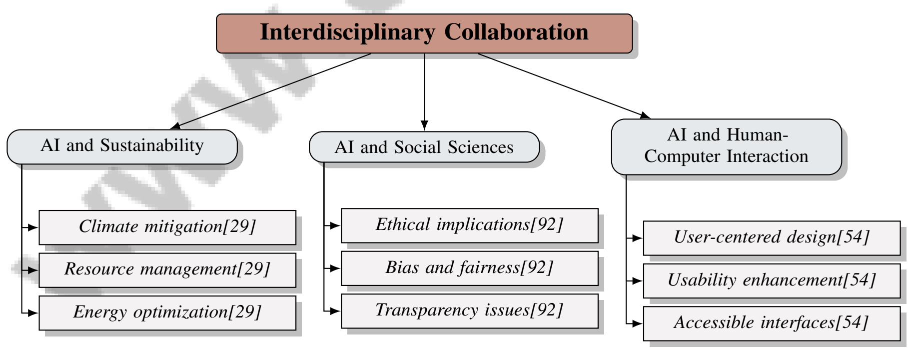


Figure 8: This figure illustrates the interdisciplinary collaboration in AI, highlighting its integration with sustainability, social sciences, and human-computer interaction to address complex challenges and enhance usability and ethical standards.

7.3 Integration of AI with Sustainability

Integrating AI with sustainability goals offers a transformative opportunity to address environmental and social challenges through technological innovation. AI technologies, particularly machine learning (ML) and deep learning, optimize processes, enhance resource efficiency, and reduce environmental impact across sectors. Developing energy-efficient algorithms and hardware significantly decreases the carbon footprint associated with AI model training and deployment [29]. By prioritizing

AI-generated, for reference only.

computational efficiency, researchers align AI advancements with sustainability objectives, promoting environmentally responsible innovation.

In the renewable energy sector, AI-driven predictive models improve energy demand forecasts, facilitating efficient energy system management [26]. These models optimize renewable energy integration into existing grids, reducing fossil fuel reliance and contributing to sustainable energy infrastructure. AI technologies enhance hybrid electric vehicles (HEVs) operational efficiency by optimizing energy usage and minimizing emissions, aligning with broader sustainability goals in transportation [75].

Leveraging AI for environmental monitoring and conservation efforts is another critical strategy. AI-powered systems analyze large datasets to identify patterns and trends in environmental changes, enabling timely interventions and informed decision-making. This capability is valuable in biodiversity conservation, supporting efforts to protect endangered species and ecosystems [20]. Integrating AI with environmental science allows stakeholders to develop targeted strategies for preserving natural resources and promoting ecological sustainability.

Adopting transparent and explainable AI frameworks that build trust among stakeholders ensures the ethical integration of AI with sustainability goals. Providing clear insights into AI decision-making processes facilitates collaboration between technologists, policymakers, and communities, aligning AI-driven solutions with societal values and priorities [92].

Integrating AI with sustainability goals requires a comprehensive strategy embracing technological innovations—such as machine learning and edge computing—while incorporating ethical frameworks and interdisciplinary collaboration. This approach addresses complex challenges outlined in the United Nations Sustainable Development Goals (SDGs), ensuring AI applications, from environmental monitoring to disaster response, are effective and responsible. Developing standardized data documentation practices, like Croissant-RAI, is critical for enhancing AI datasets' transparency and trustworthiness, mitigating biases, and fostering equitable advancements in the field [39, 18, 81, 29]. Harnessing AI's potential to drive sustainable development enables researchers and practitioners to contribute to a more equitable and environmentally conscious future.

8 Challenges and Future Directions

Addressing the integration of artificial intelligence (AI) across various domains demands a focus on data privacy and ethical considerations, establishing a framework for responsibly implementing AI technologies amid prevalent ethical dilemmas and privacy concerns. Advanced algorithms for personal data protection, such as differential privacy in machine learning, are crucial, alongside robust defenses against deepfakes and misinformation. By integrating detection strategies and advocating for ethical guidelines and public awareness, this analysis navigates AI challenges while safeguarding privacy and information integrity [98, 42, 99].

8.1 Data Privacy and Ethical Considerations

AI's integration across domains necessitates a thorough examination of data privacy and ethical considerations due to their profound impact on decision-making processes. A significant challenge is the opacity of AI models, often perceived as black boxes, raising accountability and trust concerns, particularly when documentation is insufficient [22]. Structured documentation formats are essential for addressing these issues effectively. In the context of large language models (LLMs), ethical considerations are critical, as these models can perpetuate biases, necessitating robust frameworks to evaluate and mitigate biases in AI outputs [23]. Rapid AI advancements introduce additional ethical challenges, as researchers develop detection and mitigation strategies amidst overwhelming synthetic content [21].

In decentralized decision-making, such as energy markets, ethical considerations emphasize data privacy, especially when integrating multiple distributed energy resources. Preserving individual data privacy while collaboratively improving shared models enhances forecast robustness and accuracy [3]. Privacy concerns in AI interactions necessitate careful consideration to protect sensitive information and maintain user trust [14]. The complexity of real-world systems and state data availability further complicate ethical considerations in AI applications. For instance, using LLMs for automated grading raises concerns about data privacy and ethical implications of using student data [22].

Data privacy concerns are exacerbated by the extensive training data required for AI models, raising bias and sensitive information protection issues. The emergence of adversarial examples presents further risks to deep learning applications, necessitating ongoing research into robust defensive strategies [2]. Addressing these concerns requires a comprehensive strategy prioritizing regulatory compliance, enhancing model interpretability, and adopting user-centric design while leveraging advanced machine learning algorithms for effective data anonymization [98, 94]. By prioritizing these aspects, stakeholders can develop AI systems that are effective, ethically responsible, and aligned with societal values.

8.2 Technological Limitations and Scalability

The rapid advancement of AI and machine learning (ML) technologies reveals several technological limitations and scalability challenges. A significant issue is the computational cost of traditional methods, hindering widespread adoption in high-risk domains [100]. This limitation underscores the need for more efficient algorithms in resource-constrained environments. The scalability of AI systems is further constrained by the computational demands of training complex models. The Artificial Neural Twin (ANT) framework addresses challenges like data sovereignty and continual learning, emphasizing scalable AI systems that adapt to dynamic conditions without compromising performance [101].

In reinforcement learning, methods like Advantage Weighted Regression (AWR) leverage off-policy data to update policies efficiently, but reliance on training data quality affects generalizability [102]. Additionally, limited battery capacity in learning-based sensing and computing restricts continuous data sensing and transmission, compromising data freshness [103]. Multi-task learning (MTL) offers a promising approach for enhancing model performance, reducing training costs, and improving generalization, especially with pretrained foundation models [24]. However, scalability and adaptability to new environments remain challenging, often requiring extensive tuning [71].

Integrating AI into complex systems, such as multi-agent environments, presents unique challenges in determining the effectiveness of expert interventions without hindering the agent's learning ability [8]. The DRL-MOA method demonstrates superior performance in solving large-scale bi-objective TSPs, showcasing strong generalization, fast solving speed, and promising solution quality compared to classical methods [60]. Addressing technological limitations and scalability challenges in AI and ML requires a multifaceted approach emphasizing efficient data management, computational efficiency, and scalable algorithm development. By tackling diverse challenges, including effective classification of computer-generated content and enhancing interpretability, researchers and practitioners can significantly improve AI technologies' relevance and effectiveness across sectors [16, 81, 42, 94].

8.3 Challenges and Opportunities in AI Integration

AI integration across fields presents a complex landscape of challenges and opportunities, necessitating innovative approaches to harness AI's full potential. A primary challenge is the potential for unjust punishment within AI systems, where cooperators may be unfairly penalized, diminishing incentives for cooperation [73]. This issue highlights the importance of designing AI systems capable of accurately distinguishing between cooperative and non-cooperative behaviors to ensure fair outcomes. Another challenge involves integrating AI in scenarios with limited resources or absent initial labeled data, which hampers AI systems' learning and adaptation in resource-constrained environments [104]. Addressing this challenge requires developing robust algorithms capable of operating efficiently under such constraints.

The BGM framework illustrates the challenge of bridging the preference gap between retrievers and LLMs, presenting opportunities for improved performance [21]. By aligning AI systems with user preferences, this framework enhances AI-driven solutions' effectiveness and user satisfaction. Understanding how AI agents can acquire diverse and open-ended skills, akin to human cognitive capacities, remains a significant challenge [105]. This emphasizes the need for AI systems to develop adaptive learning capabilities, enabling them to tackle a wide range of tasks and environments.

Integrating AI with sustainability metrics presents challenges, particularly in developing models that explain relationships between input factors and outcomes, like fuel consumption [14]. Enhancing model interpretability and transparency is crucial for ensuring AI systems contribute positively to environmental and social objectives. A comprehensive survey of AI integration challenges reveals that many existing solutions do not simultaneously address all identified challenges, underscoring the need for holistic approaches that consider real-world applications' multifaceted nature [74]. By developing integrated solutions addressing these challenges, researchers and practitioners can unlock new opportunities for AI to drive innovation and improve efficiency across sectors.

8.4 Interdisciplinary Collaboration and Methodological Advancements

Interdisciplinary collaboration and methodological advancements are essential for AI and ML evolution, particularly as these technologies tackle complex challenges across domains. Collaborative efforts enhance AI methodologies, improve model interpretability, and optimize system efficiency. Future research should prioritize creating comprehensive curricula for teaching ML across disciplines, exploring emerging trends in technology integration in education, and assessing ML's impact on engineering outcomes related to sustainability [18]. Integrating AI with sustainability goals necessitates refining the relearning process in reinforcement learning, leading to better adaptation and integration of AI technologies in sustainable practices [3]. This interdisciplinary approach is crucial for enhancing deep reinforcement learning algorithms' stability, improving interpretability, and exploring new applications in healthcare, finance, and autonomous systems [1].

Methodological advancements are critical in reinforcement learning, where establishing a robust theoretical foundation and exploring alternative algorithms enhance performance. Investigating distance matrices as inputs and improving solution distribution from methods like DRL-MOA significantly contribute to the field's advancement [60]. Future research could expand AI systems' emotional framework to encompass a broader range of emotional states, enriching intelligent agents' decision-making processes [2]. Promoting interdisciplinary collaboration is essential, as illustrated by initiatives like Croissant-RAI, enhancing dataset documentation and discoverability. Pursuing methodological advancements, such as developing standardized metrics for evaluating interpretability methods and creating comprehensive toolkits for AI explainability, is crucial to cater to diverse stakeholder needs [94, 106, 42, 39, 17]. By integrating diverse insights and exploring new methodological approaches, researchers create robust, transparent AI systems capable of addressing modern society's multifaceted needs.

9 Conclusion

This survey underscores the complex interdependencies among artificial intelligence (AI), machine learning (ML), deep learning (DL), reinforcement learning (RL), natural language processing (NLP), ESG ratings, and sustainability, highlighting their collective potential to drive innovation and sustainable development across various sectors. The integration of AI with ESG metrics presents promising avenues for optimizing resource management and fostering environmentally sustainable practices. Advancements in RL, particularly those enhanced by human feedback, are poised to revolutionize NLP applications, offering substantial improvements over traditional methodologies.

In the healthcare sector, the necessity of integrating explainability into AI models is pivotal for both their acceptance and efficacy. This is exemplified by methods that align closely with clinical outcomes, such as in lung cancer screening, emphasizing the importance of interpretable models that leverage diverse data sources, including social media, to improve patient care. The recognition of deep learning as a cornerstone in AI and data science further highlights its potential for future research and application development.

The survey also highlights the critical role of user-friendly explanation formats and comprehensive evaluations of explainable AI (XAI) techniques in building trust and transparency in AI systems. The concept of human-centered mechanism design, as demonstrated by innovative AI mechanisms, underscores the importance of aligning AI systems with societal values and preferences. Although neuro-symbolic AI (NeSy) holds promise for NLP, challenges persist in realizing its full potential, particularly in terms of transferability and reasoning capabilities. Nonetheless, the ongoing evolution of AI methodologies, such as the optimization of LSTM architectures, continues to produce innovative variants that surpass traditional designs, indicating a dynamic and rapidly advancing field.

References

[1] Shengbo Eben Li. Deep reinforcement learning. In *Reinforcement learning for sequential decision and optimal control*, pages 365–402. Springer, 2023.

[2] Gustavo Assunção, Miguel Castelo-Branco, and Paulo Menezes. Self-mediated exploration in artificial intelligence inspired by cognitive psychology, 2023.

[3] Avisek Naug, Marcos Quiñones-Grueiro, and Gautam Biswas. A relearning approach to reinforcement learning for control of smart buildings, 2020.

[4] Alain Andres, Esther Villar-Rodriguez, and Javier Del Ser. An evaluation study of intrinsic motivation techniques applied to reinforcement learning over hard exploration environments, 2022.

[5] Mohamed-Amine Chadi and Hajar Mousannif. Understanding reinforcement learning algorithms: The progress from basic q-learning to proximal policy optimization, 2023.

[6] Yang Cui, Yang Xu, Yang Li, Yijian Wang, and Xinpeng Zou. Deep reinforcement learning based optimal energy management of multi-energy microgrids with uncertainties, 2023.

[7] Jing Wu, Zhixin Lai, Shengjie Liu, Suiyao Chen, Ran Tao, Pan Zhao, Chuyuan Tao, Yikun Cheng, and Naira Hovakimyan. Crops: A deployable crop management system over all possible state availabilities, 2024.

[8] Adrien Bennetot, Vicky Charisi, and Natalia Díaz-Rodríguez. Should artificial agents ask for help in human-robot collaborative problem-solving?, 2020.

[9] Justin Fu, Aviral Kumar, Matthew Soh, and Sergey Levine. Diagnosing bottlenecks in deep q-learning algorithms, 2019.

[10] Kun-Peng Ning and Sheng-Jun Huang. Reinforcement learning with supervision from noisy demonstrations, 2020.

[11] Yuxi Li. Deep reinforcement learning: An overview. *arXiv preprint arXiv:1701.07274*, 2017.

[12] Francisco Durán, Silverio Martínez-Fernández, Matias Martinez, and Patricia Lago. Identifying architectural design decisions for achieving green ml serving, 2024.

[13] Jun He and Andrew L. Liu. Evaluating the impact of multiple der aggregators on wholesale energy markets: A hybrid mean field approach, 2024.

[14] Alberto Barbado and Óscar Corcho. Vehicle fuel optimization under real-world driving conditions: An explainable artificial intelligence approach, 2021.

[15] Jože M. Rožanec, Elias Montini, Vincenzo Cutrona, Dimitrios Papamartzivanos, Timotej Klemenčič, Blaž Fortuna, Dunja Mladenić, Entso Veliou, Thanassis Giannetsos, and Christos Emmanouilidis. Human in the ai loop via xai and active learning for visual inspection, 2023.

[16] Xiang Li, Zhenyu Li, Chen Shi, Yong Xu, Qing Du, Mingkui Tan, Jun Huang, and Wei Lin. Alphafin: Benchmarking financial analysis with retrieval-augmented stock-chain framework, 2024.

[17] Sean Xie, Soroush Vosoughi, and Saeed Hassanpour. Interpretation quality score for measuring the quality of interpretability methods, 2022.

[18] Andrew Schulz, Suzanne Stathatos, Cassandra Shriver, and Roxanne Moore. Utilizing online and open-source machine learning toolkits to leverage the future of sustainable engineering, 2023.

[19] Stefan Feuerriegel and Ralph Fehrer. Improving decision analytics with deep learning: The case of financial disclosures, 2018.

[20] Paul Fergus, Carl Chalmers, Steven Longmore, Serge Wich, Carmen Warmenhove, Jonathan Swart, Thuto Ngongwane, André Burger, Jonathan Ledgard, and Erik Meijaard. Empowering wildlife guardians: An equitable digital stewardship and reward system for biodiversity conservation using deep learning and 3/4g camera traps, 2023.

[21] Zixuan Ke, Weize Kong, Cheng Li, Mingyang Zhang, Qiaozhu Mei, and Michael Bendersky. Bridging the preference gap between retrievers and llms, 2024.

[22] Gloria Ashiya Katuka, Alexander Gain, and Yen-Yun Yu. Investigating automatic scoring and feedback using large language models, 2024.

[23] Guy Tennenholtz, Yinlam Chow, Chih-Wei Hsu, Lior Shani, Ethan Liang, and Craig Boutilier. Embedding-aligned language models, 2024.

[24] Jun Yu, Yutong Dai, Xiaokang Liu, Jin Huang, Yishan Shen, Ke Zhang, Rong Zhou, Eashan Adhikarla, Wenxuan Ye, Yixin Liu, Zhaoming Kong, Kai Zhang, Yilong Yin, Vinod Namboodiri, Brian D. Davison, Jason H. Moore, and Yong Chen. Unleashing the power of multi-task learning: A comprehensive survey spanning traditional, deep, and pretrained foundation model eras, 2024.

[25] Kyle Hamilton, Aparna Nayak, Bojan Božić, and Luca Longo. Is neuro-symbolic ai meeting its promise in natural language processing? a structured review, 2022.

[26] Maria Tzelepi, Charalampos Symeonidis, Paraskevi Nousi, Efstratios Kakaletsis, Theodoros Manousis, Pavlos Tosidis, Nikos Nikolaidis, and Anastasios Tefas. Deep learning for energy time-series analysis and forecasting, 2023.

[27] Jacob Paugh, Zhaoxuan Zhu, Shobhit Gupta, Marcello Canova, and Stephanie Stockar. Eco-driving control of connected and automated vehicles using neural network based rollout, 2023.

[28] Anjali Raj, Adway Mitra, and Manjira Sinha. Deep learning for slum mapping in remote sensing images: A meta-analysis and review, 2024.

[29] Emma Strubell, Ananya Ganesh, and Andrew McCallum. Energy and policy considerations for modern deep learning research. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pages 13693–13696, 2020.

[30] Poorvesh Dongre, Majid Behravan, Kunal Gupta, Mark Billinghurst, and Denis Gračanin. Integrating physiological data with large language models for empathic human-ai interaction, 2024.

[31] Yang Li, Ruinong Wang, Yuanzheng Li, Meng Zhang, and Chao Long. Wind power forecasting considering data privacy protection: A federated deep reinforcement learning approach, 2022.

[32] Adam White, Margarita Saranti, Artur d'Avila Garcez, Thomas M. H. Hope, Cathy J. Price, and Howard Bowman. Predicting recovery following stroke: deep learning, multimodal data and feature selection using explainable ai, 2023.

[33] Yu-Jia An, Sheng-Chen Bai, Lin Cheng, Xiao-Guang Li, Cheng en Wang, Xiao-Dong Han, Gang Su, Shi-Ju Ran, and Cong Wang. Intelligent diagnostic scheme for lung cancer screening with raman spectra data by tensor network machine learning, 2023.

[34] Sven Kruschel, Nico Hambauer, Sven Weinzierl, Sandra Zilker, Mathias Kraus, and Patrick Zschech. Challenging the performance-interpretability trade-off: An evaluation of interpretable machine learning models, 2024.

[35] Nham Le, Tuan Lai, Trung Bui, and Doo Soon Kim. Autonlu: An on-demand cloud-based natural language understanding system for enterprises, 2020.

[36] Zheng Wang, Bingzheng Gan, and Wei Shi. Multimodal query suggestion with multi-agent reinforcement learning from human feedback, 2024.

[37] E. A. Huerta and Zhizhen Zhao. Advances in machine and deep learning for modeling and real-time detection of multi-messenger sources, 2021.

[38] Pierre-Yves Oudeyer. Computational theories of curiosity-driven learning, 2018.

[39] Nitisha Jain, Mubashara Akhtar, Joan Giner-Miguelez, Rajat Shinde, Joaquin Vanschoren, Steffen Vogler, Sujata Goswami, Yuhan Rao, Tim Santos, Luis Oala, Michalis Karamousadakis, Manil Maskey, Pierre Marcenac, Costanza Conforti, Michael Kuchnik, Lora Aroyo, Omar Benjelloun, and Elena Simperl. A standardized machine-readable dataset documentation format for responsible ai, 2024.

[40] Andrew Slavin Ross. Training machine learning models by regularizing their explanations, 2018.

[41] Chris Roadknight, Prapa Rattadilok, and Uwe Aickelin. Teaching key machine learning principles using anti-learning datasets, 2020.

[42] Allen Lavoie and Mukkai Krishnamoorthy. Algorithmic detection of computer generated text, 2010.

[43] Yuanqing Yu, Zhefan Wang, Weizhi Ma, Shuai Wang, Chuhan Wu, Zhiqiang Guo, and Min Zhang. Steptool: Enhancing multi-step tool usage in llms through step-grained reinforcement learning, 2025.

[44] Norman Di Palo, Arunkumar Byravan, Leonard Hasenclever, Markus Wulfmeier, Nicolas Heess, and Martin Riedmiller. Towards a unified agent with foundation models, 2023.

[45] Guixuan Wen and Kaigui Wu. Building decision forest via deep reinforcement learning, 2022.

[46] Paul F Christiano, Jan Leike, Tom Brown, Miljan Martic, Shane Legg, and Dario Amodei. Deep reinforcement learning from human preferences. *Advances in neural information processing systems*, 30, 2017.

[47] Chaoyi Gu, Varuna De Silva, Corentin Artaud, and Rafael Pina. Embedding contextual information through reward shaping in multi-agent learning: A case study from google football, 2023.

[48] Ruiqi He and Falk Lieder. What are the mechanisms underlying metacognitive learning?, 2023.

[49] Randall Correll, Sean J. Weinberg, Fabio Sanches, Takanori Ide, and Takafumi Suzuki. Reinforcement learning for multi-truck vehicle routing problems, 2022.

[50] Siyi Hu, Yifan Zhong, Minquan Gao, Weixun Wang, Hao Dong, Xiaodan Liang, Zhihui Li, Xiaojun Chang, and Yaodong Yang. Marllib: A scalable and efficient multi-agent reinforcement learning library, 2023.

[51] Soma Suzuki. Multiagent-based participatory urban simulation through inverse reinforcement learning, 2017.

[52] Finn Rietz, Erik Schaffernicht, Stefan Heinrich, and Johannes Andreas Stork. Prioritized soft q-decomposition for lexicographic reinforcement learning, 2024.

[53] Matteo Turchetta, Andreas Krause, and Sebastian Trimpe. Robust model-free reinforcement learning with multi-objective bayesian optimization, 2019.

[54] Sarvesh Patil. Deep learning based natural language processing for end to end speech translation, 2018.

[55] Sean Xie, Soroush Vosoughi, and Saeed Hassanpour. Towards interpretable deep reinforcement learning models via inverse reinforcement learning, 2024.

[56] Csaba Szepesvári. *Algorithms for reinforcement learning*. Springer nature, 2022.

[57] Qiliang Chen and Babak Heydari. Adaptive network intervention for complex systems: A hierarchical graph reinforcement learning approach, 2024.

[58] Shihan Dou, Yan Liu, Haoxiang Jia, Limao Xiong, Enyu Zhou, Wei Shen, Junjie Shan, Caishuang Huang, Xiao Wang, Xiaoran Fan, Zhiheng Xi, Yuhao Zhou, Tao Ji, Rui Zheng, Qi Zhang, Xuanjing Huang, and Tao Gui. Stepcoder: Improve code generation with reinforcement learning from compiler feedback, 2024.

[59] Jiafei Lyu, Le Wan, Zongqing Lu, and Xiu Li. Off-policy rl algorithms can be sample-efficient for continuous control via sample multiple reuse, 2023.

[60] Kaiwen Li, Tao Zhang, and Rui Wang. Deep reinforcement learning for multi-objective optimization, 2020.

[61] Aniruddha Bhargava, Lalit Jain, Branislav Kveton, Ge Liu, and Subhojyoti Mukherjee. Off-policy evaluation from logged human feedback, 2024.

[62] Misha Laskin, Kimin Lee, Adam Stooke, Lerrel Pinto, Pieter Abbeel, and Aravind Srinivas. Reinforcement learning with augmented data. *Advances in neural information processing systems*, 33:19884–19895, 2020.

[63] Ryan Campbell and Junsang Yoon. Automatic curriculum learning with gradient reward signals, 2023.

[64] Pengqin Wang, Meixin Zhu, and Shaojie Shen. Environment transformer and policy optimization for model-based offline reinforcement learning, 2023.

[65] Ying Li, Zhencai Zhu, Xiaoqiang Li, Chunyu Yang, and Hao Lu. When mining electric locomotives meet reinforcement learning, 2023.

[66] Léonard Hussenot, Robert Dadashi, Matthieu Geist, and Olivier Pietquin. Show me the way: Intrinsic motivation from demonstrations, 2021.

[67] Kiran Lekkala, Eshan Bhargava, Yunhao Ge, and Laurent Itti. Evaluating pretrained models for deployable lifelong learning, 2023.

[68] Zohar Rimon, Aviv Tamar, and Gilad Adler. Meta reinforcement learning with finite training tasks – a density estimation approach, 2024.

[69] Vincent François-Lavet, Peter Henderson, Riashat Islam, Marc G Bellemare, Joelle Pineau, et al. An introduction to deep reinforcement learning. *Foundations and Trends® in Machine Learning*, 11(3-4):219–354, 2018.

[70] Antonin Raffin, Ashley Hill, Adam Gleave, Anssi Kanervisto, Maximilian Ernestus, and Noah Dormann. Stable-baselines3: Reliable reinforcement learning implementations. *Journal of machine learning research*, 22(268):1–8, 2021.

[71] Pawel Ladosz, Lilian Weng, Minwoo Kim, and Hyondong Oh. Exploration in deep reinforcement learning: A survey. *Information Fusion*, 85:1–22, 2022.

[72] Aditya Datar and Pramit Saha. The promise of analog deep learning: Recent advances, challenges and opportunities, 2024.

[73] Nayana Dasgupta and Mirco Musolesi. Investigating the impact of direct punishment on the emergence of cooperation in multi-agent reinforcement learning systems, 2024.

[74] Gabriel Dulac-Arnold, Daniel Mankowitz, and Todd Hester. Challenges of real-world reinforcement learning. *arXiv preprint arXiv:1904.12901*, 2019.

[75] Christoforos Menos-Aikateriniadis, Stavros Sykiotis, and Pavlos S. Georgilakis. Optimal scheduling of electric vehicle charging with deep reinforcement learning considering end users flexibility, 2023.

[76] Sébastien Ollivier, Sheng Li, Yue Tang, Chayanika Chaudhuri, Peipei Zhou, Xulong Tang, Jingtong Hu, and Alex K. Jones. Sustainable ai processing at the edge, 2022.

[77] Raphael Koster, Jan Balaguer, Andrea Tacchetti, Ari Weinstein, Tina Zhu, Oliver Hauser, Duncan Williams, Lucy Campbell-Gillingham, Phoebe Thacker, Matthew Botvinick, and Christopher Summerfield. Human-centered mechanism design with democratic ai, 2022.

[78] Anton Persson, Niklas Dymne, and Fernando Alonso-Fernandez. Classification of ps and abs black plastics for weee recycling applications, 2021.

[79] Shuo Wan, Jiaxun Lu, Pingyi Fan, and Khaled B. Letaief. Towards big data processing in iot: Path planning and resource management of uav base stations in mobile-edge computing system, 2019.

[80] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mane, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viegas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. Tensorflow: Large-scale machine learning on heterogeneous distributed systems, 2016.

[81] Elisa Bertino and Sujata Banerjee. Artificial intelligence at the edge, 2020.

[82] Karl Cobbe, Oleg Klimov, Chris Hesse, Taehoon Kim, and John Schulman. Quantifying generalization in reinforcement learning. In *International conference on machine learning*, pages 1282–1289. PMLR, 2019.

[83] Niful Islam, Debopom Sutradhar, Humaira Noor, Jarin Tasnim Raya, Monowara Tabassum Maisha, and Dewan Md Farid. Distinguishing human generated text from chatgpt generated text using machine learning, 2023.

[84] Joris Guerin, Raul Sena Ferreira, Kevin Delmas, and Jérémie Guiochet. Unifying evaluation of machine learning safety monitors, 2022.

[85] Jad Doughman and Wael Khreich. Gender bias in text: Labeled datasets and lexicons, 2023.

[86] Zhiqiang Wang, Yiran Pang, and Yanbin Lin. Large language models are zero-shot text classifiers, 2023.

[87] Yizhe Zhang, Jiarui Lu, and Navdeep Jaitly. Probing the multi-turn planning capabilities of llms via 20 question games, 2024.

[88] Xiangru Tang, Yuliang Liu, Zefan Cai, Yanjun Shao, Junjie Lu, Yichi Zhang, Zexuan Deng, Helan Hu, Kaikai An, Ruijun Huang, Shuzheng Si, Sheng Chen, Haozhe Zhao, Liang Chen, Yan Wang, Tianyu Liu, Zhiwei Jiang, Baobao Chang, Yin Fang, Yujia Qin, Wangchunshu Zhou, Yilun Zhao, Arman Cohan, and Mark Gerstein. Ml-bench: Evaluating large language models and agents for machine learning tasks on repository-level code, 2024.

[89] Amelia Katirai, Noa Garcia, Kazuki Ide, Yuta Nakashima, and Atsuo Kishimoto. Situating the social issues of image generation models in the model life cycle: a sociotechnical approach, 2024.

[90] Florian Scholten, Tobias R. Rebholz, and Mandy Hütter. Metacognitive myopia in large language models, 2024.

[91] Fanny Jourdan, Laurent Risser, Jean-Michel Loubes, and Nicholas Asher. Are fairness metric scores enough to assess discrimination biases in machine learning?, 2023.

[92] Satyapriya Krishna, Tessa Han, Alex Gu, Steven Wu, Shahin Jabbari, and Himabindu Lakkaraju. The disagreement problem in explainable machine learning: A practitioner's perspective, 2024.

[93] Alexander Jung and Pedro H. J. Nardelli. An information-theoretic approach to personalized explainable machine learning, 2020.

[94] Vijay Arya, Rachel K. E. Bellamy, Pin-Yu Chen, Amit Dhurandhar, Michael Hind, Samuel C. Hoffman, Stephanie Houde, Q. Vera Liao, Ronny Luss, Aleksandra Mojsilović, Sami Mourad, Pablo Pedemonte, Ramya Raghavendra, John Richards, Prasanna Sattigeri, Karthikeyan Shanmugam, Moninder Singh, Kush R. Varshney, Dennis Wei, and Yunfeng Zhang. One explanation does not fit all: A toolkit and taxonomy of ai explainability techniques, 2019.

[95] Alejandro Peña, Ignacio Serna, Aythami Morales, Julian Fierrez, Alfonso Ortega, Ainhoa Herrarte, Manuel Alcantara, and Javier Ortega-Garcia. Human-centric multimodal machine learning: Recent advances and testbed on ai-based recruitment, 2023.

[96] Alex Cummaudo, Rajesh Vasa, John Grundy, Mohamed Abdelrazek, and Andrew Cain. Losing confidence in quality: Unspoken evolution of computer vision services, 2019.

[97] Hao Wu. What is learning? a primary discussion about information and representation, 2015.

[98] Le Yang, Miao Tian, Duan Xin, Qishuo Cheng, and Jiajian Zheng. Ai-driven anonymization: Protecting personal data privacy while leveraging machine learning, 2024.

[99] Mohamed R. Shoaib, Zefan Wang, Milad Taleby Ahvanooey, and Jun Zhao. Deepfakes, misinformation, and disinformation in the era of frontier ai, generative ai, and large ai models, 2023.

[100] Mengyuan Chen, Junyu Gao, and Changsheng Xu. Revisiting essential and nonessential settings of evidential deep learning, 2024.

[101] Johannes Emmert, Ronald Mendez, Houman Mirzaalian Dastjerdi, Christopher Syben, and Andreas Maier. The artificial neural twin – process optimization and continual learning in distributed process chains, 2024.

[102] Pochun Li, Yuyang Xiao, Jinghua Yan, Xuan Li, and Xiaoye Wang. Reinforcement learning for adaptive resource scheduling in complex system environments, 2024.

[103] Sinwoong Yun, Dongsun Kim, Chanwon Park, and Jemin Lee. Learning-based sensing and computing decision for data freshness in edge computing-enabled networks, 2024.

[104] Meng Fang, Yuan Li, and Trevor Cohn. Learning how to active learn: A deep reinforcement learning approach. *arXiv preprint arXiv:1708.02383*, 2017.

[105] Eleni Nisioti and Clément Moulin-Frier. Grounding artificial intelligence in the origins of human behavior, 2020.

[106] Sicong Cao, Xiaobing Sun, Ratnadira Widyasari, David Lo, Xiaoxue Wu, Lili Bo, Jiale Zhang, Bin Li, Wei Liu, Di Wu, and Yixin Chen. A systematic literature review on explainability for machine/deep learning-based software engineering research, 2025.

Disclaimer:

SurveyX is an AI-powered system designed to automate the generation of surveys. While it aims to produce high-quality, coherent, and comprehensive surveys with accurate citations, the final output is derived from the AI's synthesis of pre-processed materials, which may contain limitations or inaccuracies. As such, the generated content should not be used for academic publication or formal submissions and must be independently reviewed and verified. The developers of SurveyX do not assume responsibility for any errors or consequences arising from the use of the generated surveys.