Novelty Assessment Report

Paper: Optimistic Task Inference for Behavior Foundation Models

PDF URL: https://openreview.net/pdf?id=m5byThUSNE

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2026-01-01

Abstract

Behavior Foundation Models (BFMs) are capable of retrieving high-performing policy for any reward function specified directly at test-time, commonly referred to as zero-shot reinforcement learning (RL). While this is a very efficient process in terms of compute, it can be less so in terms of data: as a standard assumption, BFMs require computing rewards over a non-negligible inference dataset, assuming either access to a functional form of rewards, or significant labeling efforts. To alleviate these limitations, we tackle the problem of task inference purely through interaction with the environment at test-time. We propose OpTI-BFM, an optimistic decision criterion that directly models uncertainty over reward functions and guides BFMs in data collection for task inference. Formally, we provide a regret bound for well- trained BFMs through a direct connection to upper-confidence algorithms for linear bandits. Empirically, we evaluate OpTI-BFM on established zero-shot benchmarks, and observe that it enables successor-features-based BFMs to identify and optimize an unseen reward function in a handful of episodes with minimal compute overhead.

Core Task Landscape

This paper addresses: **Task Inference for Behavior Foundation Models through Online Interaction**

A total of **31 papers** were analyzed and organized into a taxonomy with **17 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Task Inference and Reward Specification Methods**
- **Foundation Model Training Paradigms and Architectures**
- **Domain-Specific Foundation Model Applications**
- **Theoretical Frameworks and Conceptual Foundations**
- **User Behavior Modeling and Personalization Systems**

Complete Taxonomy Tree

- Task Inference for Behavior Foundation Models through Online Interaction Survey Taxonomy
- Task Inference and Reward Specification Methods
  - Optimistic and Uncertainty-Driven Task Inference ★ (1 papers)
  - [0] Optimistic Task Inference for Behavior Foundation Models (Anon et al., 2026) View paper
  - Preference-Based and Human Feedback Methods (1 papers)
  - [23] Fine-tuning Behavioral Cloning Policies with Preference-Based Reinforcement Learning (Friedrich Paul, 2025) View paper
  - Imitation Learning and Behavioral Cloning (2 papers)
  - [2] Fast imitation via behavior foundation models (M Pirotta, 2023) View paper
  - [14] Learning Models of Behavior From Demonstration and Through Interaction (Å oÅ¡iÄ, 2025) View paper
- Foundation Model Training Paradigms and Architectures
  - Multi-Task Agent Training Frameworks (2 papers)
  - [1] An interactive agent foundation model (Zane Durante, 2025) View paper
  - [21] Interactive Agent Foundation Model: A Meta-Learning Strategy (Mani Dwivedi, 2025) View paper
  - Zero-Shot and Fast Adaptation Mechanisms (1 papers)
  - [18] Fast Adaptation with Behavioral Foundation Models (Sikchi, 2025) View paper
  - Self-Play and Strategy Innovation (1 papers)
  - [12] Foundation model self-play: Open-ended strategy innovation via foundation models (Dharna, 2025) View paper
  - Whole-Body Control and Embodied Behavior Models (1 papers)
  - [17] Behavior Foundation Model for Humanoid Robots (Lu Shunlin, 2025) View paper
- Domain-Specific Foundation Model Applications
  - Autonomous Vehicles and Smart Mobility (1 papers)
  - [5] Smart mobility with agent-based foundation models: towards interactive and collaborative intelligent vehicles (Bingyi Xia, 2024) View paper
  - GUI and Digital Device Control (1 papers)
  - [6] Autoglm: Autonomous foundation agents for guis (Liu, 2024) View paper
  - Remote Sensing and Distributed Perception (1 papers)
  - [7] RS-DFM: A Remote Sensing Distributed Foundation Model for Diverse Downstream Tasks (Wang Zhechao, 2024) View paper
  - Interactive World and Game Environment Generation (2 papers)
  - [22] Can foundation models watch, talk and guide you step by step to make a cake? (Bao Yuwei, 2023) View paper
  - [24] Matrix-Game: Interactive World Foundation Model (Zhang Yifan, 2025) View paper

- Theoretical Frameworks and Conceptual Foundations
  - Decision-Making and Interactive Agent Theory (2 papers)
  - [3] Foundation models for decision making: Problems, methods, and opportunities (Yang, 2023) View paper
  - [15] Can foundation models actively gather information in interactive environments to test hypotheses? (Ke, 2024) View paper
  - Active Inference and Probabilistic Generative Models (1 papers)
  - [4] Active Inference and Humanâ Computer Interaction (Roderick Murray Smith, 2024) View paper
- User Behavior Modeling and Personalization Systems
  - Recommender Systems with Foundation Models (2 papers)
  - [8] Towards boosting llms-driven relevance modeling with progressive retrieved behavior-augmented prompting (Chen Zeyuan, 2025) View paper
  - [10] A Survey of Foundation Model-Powered Recommender Systems: From Feature-Based, Generative to Agentic Paradigms (Huang, 2025) View paper
  - Online Learning and Educational Behavior Analysis (6 papers)
  - [9] An evaluation model based on procedural behaviors for predicting MOOC learning performance: Students' online learning behavior analytics and algorithms â¦ (Y Tong, 2023) View paper
  - [13] Foundations of educational theory for online learning (Ally, 2004) View paper
  - [16] Research on the Interactive Enhancement Method of Teaching Ideological and Political Education in Colleges and Universities Based on Machine Learning (Jiabin Qi, 2024) View paper
  - [19] Construction and Algorithm Research of Online Learning Interaction Behavior Model Based on FBM (Hui Wang, 2024) View paper
  - [25] Research on online learning behavior analysis model in big data environment (Wang Peng, 2017) View paper
  - [28] Research on Intelligent Prediction of Learning Styles Based on Multiple Algorithms (Jianqiang Chen, 2024) View paper
  - Social Network and Community Behavior Modeling (6 papers)
  - [11] Classifying User Roles in Online News Forums: A Model for User Interaction and Behavior Analysis (Felix Scholz, 2024) View paper
  - [20] Semantic Modeling of Behavior in Users of Online Communities (Maxim A. Komardin, 2025) View paper
  - [26] Structural foundation for interactions in online games (Marina E. Elutina, 2021) View paper
  - [27] Modeling Online Social Behavior with a Deep Network Learning Framework (Yifan, 2023) View paper
  - [29] Learning and Inferring User Characteristics from Online Behavior and Content (Munira Syed, 2022) View paper
  - [30] A simple generative model of collective online behavior. (James P. Gleeson, 2014) View paper
  - Consumer Behavior and Sustainability Applications (1 papers)
  - [31] Primary Research: How Artificial Intelligence Can Enhance Sustainability Performance in the Fast Fashion Industry (Kogod, n.d.) View paper

Narrative

Core task: task inference for behavior foundation models through online interaction. The field addresses how agents can discover or refine their objectives by engaging with environments and users in real time, rather than relying solely on pre-specified reward functions. The taxonomy organizes this landscape into several main branches. Task Inference and Reward Specification Methods explore techniques for eliciting goals from interaction data, including optimistic exploration strategies and uncertainty-driven approaches. Foundation Model Training Paradigms and Architectures examine how large-scale models are built and adapted, spanning self-play mechanisms, imitation from demonstrations, and fine-tuning protocols. Domain-Specific Foundation Model Applications investigate deployments in areas such as robotics, web agents, and mobility systems, while Theoretical Frameworks and Conceptual Foundations provide the mathematical and conceptual underpinnings for decision-making and learning guarantees. User Behavior Modeling and Personalization Systems focus on inferring individual preferences and characteristics to tailor agent behavior accordingly.

A particularly active line of work centers on methods that balance exploration with task discovery: Optimistic Task Inference[0] exemplifies uncertainty-driven strategies that guide agents toward informative interactions, contrasting with more passive imitation approaches like Fast Imitation[2] or behavioral cloning schemes such as Fine-tuning Behavioral Cloning[23]. Meanwhile, Interactive Agent Foundation[1] and Interactive Agent Meta-Learning[21] emphasize meta-learning and rapid adaptation through online feedback, highlighting trade-offs between sample efficiency and generalization across diverse tasks. Another strand, represented by Foundation Models Decision Making[3] and Foundation Model Self-Play[12], investigates how large pre-trained models can bootstrap their own training signals or engage in self-improvement loops. Optimistic Task Inference[0] sits naturally within the uncertainty-driven cluster, sharing conceptual ground with Active Inference HCI[4] in its emphasis on proactive information gathering, yet differing in its focus on optimistic bounds rather than purely Bayesian inference. This positioning underscores ongoing questions about how best to integrate exploration incentives with foundation model architectures.

Related Works in Same Category

No sibling papers were found in the same taxonomy leaf. A taxonomy-subtopic-level comparison will be produced instead.

Taxonomy-Level Summary

The original leaf focuses on active, uncertainty-driven online exploration to infer tasks without explicit rewards or demonstrations. Its siblings represent alternative paradigms: one learns from passive expert demonstrations (imitation/behavioral cloning), while the other incorporates human preferences or feedback signals to refine task understanding. Together, these subtopics span a spectrum from passive observation to active exploration to interactive refinement in task inference for behavior foundation models.

Similarities: - All three subtopics address task inference without requiring fully specified reward functions upfront - Each approach aims to enable foundation models to understand and execute tasks with reduced manual specification - All methods can be viewed as complementary strategies for grounding behavior models in task-relevant objectives

Differences: - Optimistic and Uncertainty-Driven Task Inference relies on active online interaction and exploration driven by uncertainty, whereas Imitation Learning uses passive observation of expert demonstrations - Preference-Based Methods require human-in-the-loop feedback signals during or after training, while the original leaf operates autonomously through exploration - Imitation Learning typically requires extensive labeled datasets or demonstrations (explicitly excluded from the original leaf), while Optimistic approaches minimize interaction requirements - The original leaf emphasizes minimal interaction and uncertainty quantification, contrasting with the data-intensive nature of imitation learning and the iterative feedback cycles of preference-based methods

Suggested Search Directions: - Hybrid methods combining uncertainty-driven exploration with sparse demonstrations or preference feedback - Meta-learning approaches that enable rapid task inference across the spectrum from passive to active learning - Comparative studies on sample efficiency: minimal online interaction vs. demonstration requirements vs. human feedback iterations

Sibling Subtopics

- **Imitation Learning and Behavioral Cloning** (leaves: 1, papers: 2)

- Scope: Methods learning task-relevant behaviors directly from expert demonstrations through supervised or inverse reinforcement learning approaches.
- Exclude: Excludes methods requiring online interaction for task inference; see Optimistic and Uncertainty-Driven Task Inference.
- **Preference-Based and Human Feedback Methods** (leaves: 1, papers: 1)
- Scope: Techniques leveraging human preferences or feedback to refine policies and infer task objectives during or after initial training.
- Exclude: Excludes pure imitation from demonstrations without preference signals; see Imitation Learning and Behavioral Cloning.

Contributions Analysis

Overall novelty summary. The paper introduces OpTI-BFM, an optimistic decision criterion enabling behavior foundation models to infer task objectives through minimal online interaction rather than requiring labeled datasets or explicit reward functions. Within the taxonomy, it occupies the 'Optimistic and Uncertainty-Driven Task Inference' leaf under 'Task Inference and Reward Specification Methods'. Notably, this leaf contains only the original paper itself—no sibling papers—indicating a relatively sparse research direction within the broader field of 31 surveyed papers across multiple branches.

The taxonomy reveals neighboring approaches in sibling leaves: 'Preference-Based and Human Feedback Methods' (1 paper) and 'Imitation Learning and Behavioral Cloning' (2 papers). These alternatives address task specification through human preferences or expert demonstrations rather than autonomous exploration. The scope notes clarify that OpTI-BFM's online uncertainty-driven approach explicitly excludes offline demonstration methods and extensive labeling efforts, positioning it as a distinct paradigm. Related work on 'Zero-Shot and Fast Adaptation Mechanisms' (1 paper) shares the goal of rapid task adaptation but differs in requiring pre-learned embeddings rather than online interaction.

Among 30 candidates examined through semantic search, none provided clear refutation for any of the three core contributions: the OpTI-BFM algorithm (10 candidates examined), the regret bound via linear bandit connection (10 candidates), and the online task inference framework (10 candidates). This limited search scope suggests the specific combination of optimistic exploration, successor features, and formal regret guarantees for BFMs may represent a novel synthesis. However, the absence of refutable prior work reflects the search scale rather than exhaustive coverage of related bandit or meta-learning literature.

Given the sparse taxonomy leaf and limited search scope, the work appears to occupy a relatively unexplored niche within behavior foundation models. The formal connection to upper-confidence bandit algorithms and the focus on data-efficient task inference through interaction distinguish it from neighboring preference-based or imitation-based methods. However, the analysis covers top-30 semantic matches and does not capture potential overlaps in broader reinforcement learning or active learning communities outside the foundation model framing.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: OpTI-BFM: Optimistic Task Inference for Behavior Foundation Models

Description: The authors introduce OpTI-BFM, a method that enables task inference through active interaction with the environment at test-time rather than requiring labeled offline datasets. It uses optimistic decision-making to guide data collection by modeling uncertainty over reward functions through confidence ellipsoids.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Optimistic curiosity exploration and conservative exploitation with linear reward shaping

URL: View paper

Brief Assessment

Optimistic Curiosity Exploration[35] focuses on reward shaping through linear transformations and Q-function initialization in value-based DRL, not on task inference for behavior foundation models through active interaction with uncertainty modeling over reward functions.

---

2. Exploring pessimism and optimism dynamics in deep reinforcement learning

URL: View paper

Brief Assessment

Pessimism Optimism Dynamics[33] focuses on balancing pessimism and optimism in off-policy actor-critic algorithms for continuous control, not on task inference for behavior foundation models using uncertainty over reward functions.

---

3. Exploit reward shifting in value-based deep-rl: Optimistic curiosity-based exploration and conservative exploitation via linear reward shaping

URL: View paper

Brief Assessment

Exploit Reward Shifting[40] focuses on reward shaping in value-based deep RL through linear transformations affecting Q-function initialization, not on task inference for behavior foundation models through active environment interaction.

---

4. Randomized Exploration for Reinforcement Learning with General Value Function Approximation

URL: View paper

Brief Assessment

Randomized Exploration[34] focuses on general value function approximation in model-free RL with exploration via data perturbation, not on task inference for behavior foundation models using uncertainty over reward functions in a zero-shot RL setting.

---

5. Uncertainty Based Exploration in Reinforcement Learning

URL: View paper

Brief Assessment

Uncertainty Based Exploration[39] focuses on Bayesian Deep Q-Networks for exploration in standard RL environments (Atari, Cart Pole, Deep Sea), not on task inference for behavior foundation models using optimistic decision-making over reward functions.

---

6. MetaCARD: meta-reinforcement learning with task uncertainty feedback via decoupled context-aware reward and dynamics components

URL: View paper

Brief Assessment

MetaCARD[36] focuses on meta-RL with decoupled context-aware reward and dynamics components for task inference, not on optimistic exploration using confidence ellipsoids over reward functions for BFMs.

---

7. Bayesian optimistic optimization: Optimistic exploration for model-based reinforcement learning

URL: View paper

Brief Assessment

Bayesian Optimistic Optimization[41] focuses on model-based RL with optimistic exploration for learning environment dynamics in MDPs, not on task inference for behavior foundation models using pre-trained successor features.

---

8. Reinforcement learning under uncertainty: Expected versus unexpected uncertainty and state versus reward uncertainty

URL: View paper

Brief Assessment

Expected Unexpected Uncertainty[37] focuses on reinforcement learning under perceptual/state uncertainty versus reward uncertainty in psychophysical tasks, not on optimistic exploration methods for task inference in behavior foundation models using confidence ellipsoids over reward functions.

---

9. Beyond optimism: Exploration with partially observable rewards

URL: View paper

Brief Assessment

Beyond Optimism Exploration[32] focuses on optimistic exploration in partially observable reward settings with tabular MDPs, while the original paper addresses task inference for behavior foundation models using successor features in continuous control environments. The technical approaches and problem settings are fundamentally different.

---

10. Improved Regret Bound for Safe Reinforcement Learning via Tighter Cost Pessimism and Reward Optimism

URL: View paper

Brief Assessment

Safe Reinforcement Pessimism[38] addresses safe RL in tabular constrained MDPs with cost constraints, while the original paper focuses on task inference for behavior foundation models using optimistic exploration over reward functions. These are fundamentally different problem settings and methodological approaches.

Contribution 2: Regret bound for well-trained BFMs via linear bandit connection

Description: The authors establish theoretical guarantees by connecting the task inference problem to linear contextual bandits, proving that OpTI-BFM achieves sublinear regret when the underlying BFM is well-trained and certain assumptions hold.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Can large language models explore in-context?

URL: View paper

Brief Assessment

LLMs Explore In-Context[53] focuses on multi-armed bandits with LLMs as agents, not behavior foundation models with successor features. The candidate establishes regret bounds for LLM-based bandit algorithms, while the original work connects task inference in BFMs to linear contextual bandits—fundamentally different problem settings and model architectures.

---

2. Transformers as decision makers: Provable in-context reinforcement learning via supervised pretraining

URL: View paper

Brief Assessment

Transformers Decision Makers[57] analyzes supervised pretraining for in-context RL with transformers approximating LinUCB/Thompson sampling for linear bandits, while the original paper studies online task inference for behavior foundation models using successor features with UCB-based exploration.

---

3. Jump starting bandits with llm-generated prior knowledge

URL: View paper

Brief Assessment

Jump Starting Bandits[60] focuses on using LLMs to initialize contextual multi-armed bandits for recommendation systems, not on establishing regret bounds for behavior foundation models through connections to linear contextual bandits in reinforcement learning settings.

---

4. Convergence-aware online model selection with time-increasing bandits

URL: View paper

Brief Assessment

Convergence-aware Model Selection[52] addresses online model selection with time-increasing bandits, focusing on model performance convergence during training/finetuning. The original paper connects task inference in BFMs to linear contextual bandits for regret bounds, which is a fundamentally different problem setting and technical approach.

---

5. Sequential query prediction based on multi-armed bandits with ensemble of transformer experts and immediate feedback

URL: View paper

Brief Assessment

Sequential Query Prediction[59] addresses query prediction in interactive data exploration using multi-armed bandits with transformer experts, not regret bounds for behavior foundation models connected to linear contextual bandits in reinforcement learning settings.

---

6. Understanding the training and generalization of pretrained transformer for sequential decision making

URL: View paper

Brief Assessment

Pretrained Transformer Sequential[56] focuses on supervised pre-trained transformers for sequential decision-making without transition probabilities (bandits, pricing), not behavior foundation models with successor features in MDPs.

---

7. Supervised pretraining can learn in-context reinforcement learning

URL: View paper

Brief Assessment

Supervised Pretraining In-Context[51] focuses on in-context learning for transformers in RL settings, establishing regret bounds through posterior sampling equivalence. The original paper connects BFMs to linear contextual bandits for task inference, a different technical approach and problem setting.

8. Llms-augmented contextual bandit

URL: View paper

Brief Assessment

LLMs-augmented Contextual Bandit[54] focuses on using LLMs as context encoders for traditional contextual bandits with action selection, not on behavior foundation models or task inference in zero-shot RL settings.

9. Pretraining decision transformers with reward prediction for in-context multi-task structured bandit learning

URL: View paper

Brief Assessment

Pretraining Decision Transformers[58] focuses on multi-task structured bandits using transformers for in-context learning, not on establishing regret bounds through linear contextual bandit connections for behavior foundation models.

10. Contextual Bandit Optimization with Pre-Trained Neural Networks

URL: View paper

Brief Assessment

Contextual Bandit Pretrained[55] focuses on neural network contextual bandits with pre-trained representation networks, not behavior foundation models or successor features. The technical setting and problem formulation differ fundamentally from the original paper's BFM framework.

Contribution 3: Online task inference framework for BFMs

Description: The authors propose a new framework where task inference occurs online during deployment by actively collecting data, removing the need for pre-training dataset access and reducing labeling requirements compared to standard offline inference pipelines.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Adaptive stream processing on edge devices through active inference

URL: View paper

Brief Assessment

Active Inference Edge[42] focuses on adaptive stream processing on edge devices using active inference principles for SLO fulfillment, not on online task inference for behavior foundation models through active data collection during deployment.

2. CODE: Fast and Accurate Inference for Compact Distributed IoT Data Collection

URL: View paper

Brief Assessment

CODE[43] focuses on data completion tasks in distributed IoT systems with offline learning and online deployment stages, not on behavior foundation models or reinforcement learning task inference through active data collection.

3. Active Data Curation Effectively Distills Large-Scale Multimodal Models

URL: View paper

Brief Assessment

Active Data Curation[47] focuses on data selection for knowledge distillation in multimodal models, not on online task inference for behavior foundation models during deployment.

4. Runtime performance anomaly diagnosis in production hpc systems using active learning

URL: View paper

Brief Assessment

Runtime Anomaly Diagnosis[46] focuses on diagnosing performance anomalies in HPC systems using active learning on telemetry data, not on task inference for behavior foundation models during deployment in reinforcement learning settings.

5. Bridging Active Exploration and Uncertainty-Aware Deployment Using Probabilistic Ensemble Neural Network Dynamics

URL: View paper

Brief Assessment

Probabilistic Ensemble Dynamics[49] focuses on model-based RL with dynamics learning for robotics control, not on behavior foundation models or task inference through active data collection during deployment.

6. Collect & infer-a fresh look at data-efficient reinforcement learning

URL: View paper

Brief Assessment

Collect and Infer[48] proposes a general paradigm for separating data collection and inference in RL, but does not specifically address task inference for behavior foundation models or the online inference of reward functions during deployment that is central to the original paper's contribution.

7. Task Assignment Scheme Designed for Online Urban Sensing Based on Sparse Mobile Crowdsensing

URL: View paper

Brief Assessment

Sparse Mobile Crowdsensing[45] addresses urban sensing task assignment with mobility constraints and data inference, not online task inference for behavior foundation models through active data collection during deployment.

8. Active Inference and Humanâ€"Computer Interaction

URL: View paper

Brief Assessment

Active Inference HCI[4] focuses on human-computer interaction modeling using active inference theory for interface design and adaptation, not on behavior foundation models or reinforcement learning task inference during deployment.

9. Multi-modal active perception for information gathering in science missions

URL: View paper

Brief Assessment

Multi-modal Active Perception[50] addresses active perception for robotic science missions using Bayesian networks and information gain maximization. This is fundamentally different from the ORIGINAL paper's online task inference for behavior foundation models in reinforcement learning settings.

10. Active Learning of Runtime Monitors Under Uncertainty

URL: View paper

Brief Assessment

Active Learning Monitors[44] addresses runtime monitor learning for cyber-physical systems with sensor uncertainty, not task inference for behavior foundation models in reinforcement learning. The domains and technical approaches are fundamentally different.

Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

References

- [0] Optimistic Task Inference for Behavior Foundation Models View paper
- [1] An interactive agent foundation model View paper
- [2] Fast imitation via behavior foundation models View paper
- [3] Foundation models for decision making: Problems, methods, and opportunities View paper
- [4] Active Inference and Humanâ Computer Interaction View paper
- [5] Smart mobility with agent-based foundation models: towards interactive and collaborative intelligent vehicles View paper
- [6] Autoglm: Autonomous foundation agents for guis View paper
- [7] RS-DFM: A Remote Sensing Distributed Foundation Model for Diverse Downstream Tasks View paper
- [8] Towards boosting llms-driven relevance modeling with progressive retrieved behavior-augmented prompting View paper
- [9] An evaluation model based on procedural behaviors for predicting MOOC learning performance: Students' online learning behavior analytics and algorithms â¦ View paper
- [10] A Survey of Foundation Model-Powered Recommender Systems: From Feature-Based, Generative to Agentic Paradigms View paper
- [11] Classifying User Roles in Online News Forums: A Model for User Interaction and Behavior Analysis View paper
- [12] Foundation model self-play: Open-ended strategy innovation via foundation models View paper
- [13] Foundations of educational theory for online learning View paper
- [14] Learning Models of Behavior From Demonstration and Through Interaction View paper
- [15] Can foundation models actively gather information in interactive environments to test hypotheses? View paper
- [16] Research on the Interactive Enhancement Method of Teaching Ideological and Political Education in Colleges and Universities Based on Machine Learning View paper
- [17] Behavior Foundation Model for Humanoid Robots View paper
- [18] Fast Adaptation with Behavioral Foundation Models View paper
- [19] Construction and Algorithm Research of Online Learning Interaction Behavior Model Based on FBM View paper
- [20] Semantic Modeling of Behavior in Users of Online Communities View paper
- [21] Interactive Agent Foundation Model: A Meta-Learning Strategy View paper
- [22] Can foundation models watch, talk and guide you step by step to make a cake? View paper
- [23] Fine-tuning Behavioral Cloning Policies with Preference-Based Reinforcement Learning View paper
- [24] Matrix-Game: Interactive World Foundation Model View paper
- [25] Research on online learning behavior analysis model in big data environment View paper
- [26] Structural foundation for interactions in online games View paper
- [27] Modeling Online Social Behavior with a Deep Network Learning Framework View paper
- [28] Research on Intelligent Prediction of Learning Styles Based on Multiple Algorithms View paper
- [29] Learning and Inferring User Characteristics from Online Behavior and Content View paper
- [30] A simple generative model of collective online behavior. View paper
- [31] Primary Research: How Artificial Intelligence Can Enhance Sustainability Performance in the Fast Fashion Industry View paper
- [32] Beyond optimism: Exploration with partially observable rewards View paper
- [33] Exploring pessimism and optimism dynamics in deep reinforcement learning View paper
- [34] Randomized Exploration for Reinforcement Learning with General Value Function Approximation View paper
- [35] Optimistic curiosity exploration and conservative exploitation with linear reward shaping View paper
- [36] MetaCARD: meta-reinforcement learning with task uncertainty feedback via decoupled context-aware reward and dynamics components View paper
- [37] Reinforcement learning under uncertainty: Expected versus unexpected uncertainty and state versus reward uncertainty View paper
- [38] Improved Regret Bound for Safe Reinforcement Learning via Tighter Cost Pessimism and Reward Optimism View paper
- [39] Uncertainty Based Exploration in Reinforcement Learning View paper
- [40] Exploit reward shifting in value-based deep-rl: Optimistic curiosity-based exploration and conservative exploitation via linear reward shaping View paper
- [41] Bayesian optimistic optimization: Optimistic exploration for model-based reinforcement learning View paper
- [42] Adaptive stream processing on edge devices through active inference View paper
- [43] CODE: Fast and Accurate Inference for Compact Distributed IoT Data Collection View paper
- [44] Active Learning of Runtime Monitors Under Uncertainty View paper
- [45] Task Assignment Scheme Designed for Online Urban Sensing Based on Sparse Mobile Crowdsensing View paper

- [46] Runtime performance anomaly diagnosis in production hpc systems using active learning View paper
- [47] Active Data Curation Effectively Distills Large-Scale Multimodal Models View paper
- [48] Collect & infer-a fresh look at data-efficient reinforcement learning View paper
- [49] Bridging Active Exploration and Uncertainty-Aware Deployment Using Probabilistic Ensemble Neural Network Dynamics View paper
- [50] Multi-modal active perception for information gathering in science missions View paper
- [51] Supervised pretraining can learn in-context reinforcement learning View paper
- [52] Convergence-aware online model selection with time-increasing bandits View paper
- [53] Can large language models explore in-context? View paper
- [54] Llms-augmented contextual bandit View paper
- [55] Contextual Bandit Optimization with Pre-Trained Neural Networks View paper
- [56] Understanding the training and generalization of pretrained transformer for sequential decision making View paper
- [57] Transformers as decision makers: Provable in-context reinforcement learning via supervised pretraining View paper
- [58] Pretraining decision transformers with reward prediction for in-context multi-task structured bandit learning View paper
- [59] Sequential query prediction based on multi-armed bandits with ensemble of transformer experts and immediate feedback View paper
- [60] Jump starting bandits with llm-generated prior knowledge View paper