arXiv Articles Digest for 2025 July

Neruthes

2025-07-31

Abstract

This is a simple compilation of arXiv entries about LLM for 2025 July.

Full List

2507.16835
Evaluating Speech-to-Text x LLM x Text-to-Speech Combinations for AI Interview Systems

2507.16841
AquaChat: An LLM-Guided ROV Framework for Adaptive Inspection of Aquaculture Net Pens

2507.16852
SynthCTI: LLM-Driven Synthetic CTI Generation to enhance MITRE Technique Mapping

2507.16860
Weak Links in LinkedIn: Enhancing Fake Profile Detection in the Age of LLMs

2507.16951
Harnessing RLHF for Robust Unanswerability Recognition and Trustworthy Response Generation in LLMs

2507.16969
LLM4MEA: Data-free Model Extraction Attacks on Sequential Recommenders via Large Language Models

2507.16974
Leveraging Synthetic Data for Question Answering with Multilingual LLMs in the Agricultural Domain

2507.16989
Obscured but Not Erased: Evaluating Nationality Bias in LLMs via Name-Based Bias Benchmarks

2507.17015
Can External Validation Tools Improve Annotation Quality for LLM-as-a-Judge?

2507.17016
Causal Graph Fuzzy LLMs: A First Introduction and Applications in Time Series Forecasting

2507.17061
Parallelism Meets Adaptiveness: Scalable Documents Understanding in Multi-Agent LLM Systems

2507.17075
LoRA is All You Need for Safety Alignment of Reasoning LLMs

2507.17080
VL-CLIP: Enhancing Multimodal Recommendations via Visual Grounding and LLM-Augmented CLIP Embeddings

2507.17120
BucketServe: Bucket-Based Dynamic Batching for Smart and Efficient LLM Inference Serving

2507.17133
BrownoutServe: SLO-Aware Inference Serving under Bursty Workloads for MoE-based LLMs

2507.17134
Resilient Multi-Agent Negotiation for Medical Supply Chains:Integrating LLMs and Blockchain for Transparent Coordination

2507.17147
CogDual: Enhancing Dual Cognition of LLMs via Reinforcement Learning with Implicit Rule-Based Rewards

2507.17165
Can LLMs Write CI? A Study on Automatic Generation of GitHub Actions Configurations

2507.17168
Improving LLMs' Generalized Reasoning Abilities by Graph Problems

2507.17178
SKA-Bench: A Fine-Grained Benchmark for Evaluating Structured Knowledge Understanding of LLMs

2507.17188
LLM Meets the Sky: Heuristic Multi-Agent Reinforcement Learning for Secure Heterogeneous UAV Networks

2507.19525
MMCircuitEval: A Comprehensive Multimodal Circuit-Focused Benchmark for Evaluating LLMs

2507.19537
Mind the Language Gap in Digital Humanities: LLM-Aided Translation of SKOS Thesauri

2507.19549
AccessGuru: Leveraging LLMs to Detect and Correct Web Accessibility Violations in HTML Code

2507.19562
PennyCoder: Efficient Domain-Specific LLMs for PennyLane-Based Quantum Code Generation

2507.19570
MCP4EDA: LLM-Powered Model Context Protocol RTL-to-GDSII Automation with Backend Aware Synthesis Optimization

2507.19608
DeltaLLM: A Training-Free Framework Exploiting Temporal Sparsity for Efficient Edge LLM Inference

2507.19643
Can You Share Your Story? Modeling Clients' Metacognition and Openness for LLM Therapist Evaluation

2507.19747
TokenBlowUp: Resolving Representational Singularities in LLM Token Spaces via Monoidal Transformations

2507.19749
Can LLMs Solve ASP Problems? Insights from a Benchmarking Study (Extended Version)

2507.19823
HCAttention: Extreme KV Cache Compression via Heterogeneous Attention Computing for LLMs

2507.19845
MegatronApp: Efficient and Comprehensive Management on Distributed LLM Training

2507.19855
Inducing Causal World Models in LLMs for Zero-Shot Physical Reasoning

2507.19899
A Gold Standard Dataset and Evaluation Framework for Depression Detection and Explanation in Social Media using LLMs

2507.19939
LLMControl: Grounded Control of Text-to-Image Diffusion-based Synthesis with Multimodal LLMs

2507.19956
Predicting Brain Responses To Natural Movies With Multimodal LLMs

2507.19980
Exploring LLM Autoscoring Reliability in Large-Scale Writing Assessments Using Generalizability Theory

2507.20059
RAG in the Wild: On the (In)effectiveness of LLMs with Mixture-of-Knowledge Retrieval Augmentation

2507.20066
Studying Disinformation Narratives on Social Media with LLMs and Semantic Similarity

2507.20067
PITA: Preference-Guided Inference-Time Alignment for LLM Post-Training

2507.20147
Integrating LLM-Derived Multi-Semantic Intent into Graph Model for Session-based Recommendation

2507.20152
Goal Alignment in LLM-Based User Simulators for Conversational AI

2507.20208
IQ Test for LLMs: An Evaluation Framework for Uncovering Core Skills in LLMs

2507.20215
MLC-Agent: Cognitive Model based on Memory-Learning Collaboration in LLM Empowered Agent Simulation Environment

2507.20278
MoL-RL: Distilling Multi-Step Environmental Feedback into LLMs for Feedback-Independent Reasoning

2507.20300
Talking-to-Build: How LLM-Assisted Interface Shapes Player Performance and Experience in Minecraft

2507.20352
RMTBench: Benchmarking LLMs Through Multi-Turn User-Centric Role-Playing

2507.20474
MountainLion: A Multi-Modal LLM-Based Agent System for Interpretable and Adaptive Financial Trading

2507.20509
LLMs-guided adaptive compensator: Bringing Adaptivity to Automatic Control Systems with Large Language Models

2507.20511
Beyond Class Tokens: LLM-guided Dominant Property Mining for Few-shot Classification

2507.20527
SAND-Math: Using LLMs to Generate Novel, Difficult and Useful Mathematics Questions and Answers

2507.20541
MeLA: A Metacognitive LLM-Driven Architecture for Automatic Heuristic Design

2507.20655
CoGrader: Transforming Instructors' Assessment of Project Reports through Collaborative LLM Integration

2507.20666
MIMII-Agent: Leveraging LLMs with Function Calling for Relative Evaluation of Anomalous Sound Detection

2507.20674
LLM-Based Repair of Static Nullability Errors

2507.20774
evalSmarT: An LLM-Based Framework for Evaluating Smart Contract Generated Comments

2507.20849
Latent Inter-User Difference Modeling for LLM Personalization

2507.20870
A Human-in-the-loop Approach to Robot Action Replanning through LLM Common-Sense Reasoning

2507.20957
Your AI, Not Your View: The Bias of LLMs in Investment Analysis

2507.20977
Repairing vulnerabilities without invisible hands. A differentiated replication study on LLMs

2507.20999
LoRA-PAR: A Flexible Dual-System LoRA Partitioning Approach to Efficient LLM Fine-Tuning

2507.21017
MIRAGE-Bench: LLM Agent is Hallucinating and Where to Find Them

2507.21028
Multi-Agent-as-Judge: Aligning LLM-Agent-Based Automated Evaluation with Multi-Dimensional Human Evaluation