Yishan Jiang

Email: yishanjiang05@gmail.com — Phone: +86 13121910068

Education

Beijing University of Technology (BJUT) Beijing, China
B.S. in Computer Science and Technology 2022–2026
Core Courses: Data Structures and Algorithms, Computer Architecture, Operating Systems, Deep Learning, Distributed Systems, Database Systems

Publications

[**1**] Paper on mitigating topic dilution in LLM-based clustering (co-first, under review at KDD 2026)

[**2**] DualSG (ACM MM 2025) — Contributed to ablation experiments and visualization `https://arxiv.org/abs/2507.21830`

Research Experience

Intern Researcher Oct 2023–Present
Chinese Academy of Sciences (CAS), Beijing, China

Core Research: Mitigating Topic Dilution in LLMs (*KDD 2026 Submission*)

- **Problem Definition:** Formalized the *Topic Dilution* phenomenon in LLM-based clustering. Discovered that LLMs tend to generate factually correct but non-discriminative hypernyms due to *contextual isolation* and a lack of global topological signals.
- **Methodology:** Developed a framework bridging global statistical analysis with LLM inference. Executed precise concept ablation in the embedding space to extract discriminative anchors, injecting them as contextual constraints to guide and bound the model's semantic generation.
- **Results:** Reduced generated semantic entropy by approximately 75% and improved the P@5 metric from 0.37 to 0.71 in downstream retrieval tasks.

Mechanistic Interpretability of Topic Dilution (*Early Exploration*)

- Using **TransformerLens** to explore the internal mechanisms behind the Topic Dilution phenomenon, with initial experiments examining how LLMs process and generate non-discriminative hypernymic responses.

AEbench: Human-Aligned Motion Amplitude Benchmark for Video Generation

- **Motivation:** Identified a systematic bias in existing evaluation frameworks: the absence of explicit motion amplitude metrics causes benchmarks to favor static or slow-motion outputs. Existing solutions — LLM+MLLM pipelines (high scoring variance) and LLM+optical-flow pipelines (background-motion and noise sensitivity) — each exhibit distinct failure modes.
- **Design:** Proposed a dual-perspective evaluation architecture combining a *Text Disentangler* (LLM) for per-object semantic motion criteria extraction and a *Video Disentangler* (CV) for feature-level motion analysis, integrating both views through a confidence-weighted multi-round scoring mechanism.

- **Implementation:** Built a multi-agent LLM/MLLM pipeline with 4 specialized agents and multi-provider support; CV module integrates RAFT optical flow, UNISAL saliency weighting, and PAS object detection.

Scientific Impact Evaluation & Tech-Talent Ranking System

- **Data Infrastructure:** Processed over 160 million academic records from the OpenAlex dataset using PostgreSQL to construct practitioner profiles; designed efficient custom indexes to overcome physical memory limits during large-scale relational queries.
- **Semantic Feature Extraction:** Deployed the llm2vec model to generate semantic vector representations for academic entities, supporting downstream classification and ranking algorithms.

AIGCBench: Multimodal Video Generation Evaluation Nov 2023–Aug 2024

- Contributed to an evaluation benchmark for text-to-video generation models under Dr. Fanda Fan, gaining foundational exposure to diffusion model evaluation methodology and benchmark design.
- Conducted systematic literature review on benchmark design and data collection pipelines, establishing the analytical foundation for subsequent independent research.

Projects

PintOS Operating System Dec 2024–Jan 2025
Based on Stanford CS140

- Implemented priority scheduling and synchronization mechanisms in the thread module using C.
- Developed system calls for the user program module, demonstrating solid low-level memory management and system-level debugging skills.

Pet Care Robot System 2024

- Deployed a pre-trained audio classification model from HuggingFace to perform real-time pet mood inference on embedded hardware, integrating model outputs into an automated behavioral response pipeline.

Technical Skills

Machine Learning & LLM Tools: PyTorch, HuggingFace Transformers, TransformerLens, scikit-learn
Programming & Systems: Python, PostgreSQL, Linux/Bash Scripting