Yiran Jiao

Website: https://yiranjiao.github.io/ | E-mail: jiaoyiran0720@outlook.com | Tel: +86 13620467243
Address: Yunshang Community, Heiquan Road, Haidian District, Beijing, China 100192

RESEARCH INTERESTS

My research interest lies in the intersection of data science and artificial intelligence, with particular emphasis on multimodal large language models, generative modeling, and data-driven representation learning, aiming to advance robust and generalizable intelligent systems.

EDUCATION

The University of Manchester, Manchester, UK 09/2021–09/2022
Master of Science | *Health Data Science*

- Major Courses (**All Pass with Merit**) : Machine learning & ADM, Fundamental Math & Stats, Modern Information Engineering, Fundamental Mathematics and Statistic, Machine learning and Advanced Data Method, etc.
- Dissertation: Correlation between Behavioral Sensing and Ecological Momentary Assessment in Schizophrenia Patients based on Smartphone Data

Beijing Normal University - Hong Kong Baptist University United International College, HongKong) 09/2017–06/2021
Bachelor of Science | *Statistics*

- Award: 2020-2021 First-class Scholarship of the University **(TOP 3%)**
- Major Courses: (Major GPA: **3.51/4.00**) Advanced Statistics(A), Regression Analysis(A), Linear Programming and Integer Programming(A-), Data Mining(A-), Network and Transportation Models(A), Loss Model(A), Categorical Data Analysis(A), etc.

PROFESSIONAL EXPERIENCES

Kwai Technology Co., Ltd, Beijing, China 06/2025–Present
Commercial Product Department | *AI Product Manager*

Digital Human Livestreaming (Multimodal Script Generation & Automated Livestream Orchestration)

Delivered an automated livestreaming system on Kwai (China's second-largest short video platform after TikTok's local version), generating **$0.57M/d** advertising revenue and **$1.41M/d** GMV. Our system initiates a new "zero-configuration, fully managed livestream" paradigm.

- Designed a multi-agent orchestration architecture (product analysis, livestream layout, script generation, avatar selection, interaction control) via MCP, enabling cooperative task planning, multimodal content generation, and automated livestream session lifecycle management.
- Built a multimodal evaluation framework assessing text–audio–vision coherence, script semantics, and layout richness to ensure generation quality consistency across scenarios.
- Led A/B tests and applied statistical hypothesis testing to quantify advertising effect improvements from model updates. Insights directly guided model iteration direction and priorities.

BYD Co., Ltd, Shenzhen, China 12/2022–05/2025
AI Lab | *AI Product Manager*

AI Customer Service (Conversational AI and Knowledge Base Construction)

Responsible for the full product lifecycle—design, implementation, optimization—of the company's AI Customer Service ecosystem (Knowledge Base Platform, Agent Assist, Text Bot, Voice Bot, and AI Ops), achieving **80%+** recommendation accuracy | **90%+** assistant accuracy | **98%** auto-record accuracy | **49.8%** shorter service time.

- Constructed domain-specific fine-tuning datasets with tens of thousands of high-quality annotated samples per iteration, using data mining, denoising, clustering, and semantic diagnostics (intent entropy, coverage matrices, error attribution).
- Developed a customized RAG pipeline, improving knowledge QA accuracy to **90%+**, integrating multi-path retrieval, domain-term mapping, and structured knowledge parsing.
- Created a workflow-driven agent system enabling static composition of multi-step business flows with explicit tool invocation, including: CRM system queries, service-log diagnostics, order-status verification APIs, etc. Enabled autonomous root-cause analysis and guided issue resolution across complex after-sales scenarios.

RESEARCH PROJECTS

Multimodal Real-Scene Digital Human Dynamic Orchestration for Multi-Angle Live-stream Driven Generation 11/2025

Effectively addressed major production bottlenecks in digital-human livestreaming, including: single-angle visual materials, limited large-angle body motion, unnatural interaction and the sense of distance.

- Designed an ASR semantic understanding pipeline to automatically segment raw video materials into product explanation, marketing persuasion, and interactive engagement atomic clips via speech-to-text and semantic classification.
- Leveraged first–last frame video generation based on the Kling-V2 model (**TOP 1** model in video generation track released by Artificial Analysis) [1] ability to construct a dynamic arrangement and real-time driven generation framework for real-scene digital human footage.

Custom Retrieval-Augmented Generation (RAG) Framework for Enterprise Knowledge Systems 05/2024

Designed a domain-adapted RAG architecture for the New Energy Vehicles Industry's private knowledge base, achieving 90%+ answer accuracy. Addressed limitations of general RAG frameworks by enabling structured retrieval, noise reduction, long-document reasoning, and robust handling of complex, multi-format enterprise materials.

- Built and maintained a large-scale private knowledge base consisting of over **100,000** structured & unstructured documents, segmented into million-level semantic chunks via OCR recognition, hierarchical document parsing, domain-aware chunking, and metadata indexing.
- Implemented customized retrieval strategies including: (1) hybrid retrieval and context-aware re-ranking, (2) document chunking & semantic slicing optimized for long technical documents, (3) domain term mapping and entity normalization for consistency across heterogeneous documentation formats.

PROFESSIONAL SKILLS

Software: C/C++; Python; R Language; SQL; MATLAB; Tableau

[1] Ding, Y. et al., Kling-Avatar: Grounding Multimodal Instructions for Photorealistic Long-Duration Avatar Video Generation, arXiv, 2025.