

ALLYNE ZHANG

Mobile: +1 4129096972 | Email: allynez@andrew.cmu.edu | [LinkedIn](#) | [GitHub](#)

EDUCATION

Carnegie Mellon University | M.S. in Artificial Intelligence and Innovation May 2026
• GPA: 4.22/4.33, Ranked 1st in cohort
• Selected coursework: LLMs, Deep Reinforcement Learning, Multimodal ML, Deep Learning Systems, ML in Production, Search Engines

Singapore Management University | B.S. in Information Systems; Major in Business Analytics May 2024
• GPA: 4.00/4.00
• Awards: Award for Top Graduating Student in Information Systems

EXPERIENCE

Search Engines Teaching Assistant | Carnegie Mellon University, Pittsburgh, PA Aug 2025 – Dec 2025
• Serve as a teaching assistant for “Search Engines” under Prof. Jamie Callan; assist with grading and hold consultation sessions

Machine Learning Engineer Intern | ByteDance, San Jose, CA May 2025 – Aug 2025
• Worked on the Ads Targeting team under Ads Core in the Global Monetization Product and Technology department
• Developed a 7B LLM using PyTorch with supervised fine-tuning (SFT) and reinforcement learning (RLHF) to generate personalized ad targeting settings for TikTok Ads, deployed in production
• Designed robust offline evaluation frameworks and human-in-the-loop testing to reduce hallucinations and improve response quality
• Researched LLM-based approaches for improving user embeddings, including lifelong user behavior modeling

Algorithm Engineer Intern | Shopee, Singapore May 2024 – Aug 2024
• Shopee is the largest e-commerce platform in Southeast Asia, attracting 343 million monthly visitors
• Built personalized recommendation models using PyTorch for the customer service chatbot, collaborating with product and engineering teams to ship features that increased click-through rate (CTR) by over 3%
• Conducted large-scale hyperparameter tuning and optimization of attention-based models, improving training efficiency and accuracy in production-scale systems
• Analyzed A/B test data using SQL, Pandas, and plotting libraries to extract actionable insights
• Researched the application of LLMs for improving Recommendation Systems

AI Engineer Intern | Infocomm Media Development Authority, Singapore Dec 2023 – Apr 2024
• Built a summarization and question answering (QA) system for *LawNet*, Singapore’s leading legal research portal
• Improved summarization accuracy & latency for an in-house 13B LLM via multi-modal document processing, offline evaluation enhancements, benchmarking, hierarchical text chunking, and optimization techniques for long-context models

Student Researcher | Singapore Management University, Singapore Aug 2023 – Nov 2023
• Researched continuous learning for virtual reality (VR) AI agents with Professor Tony Tang in RICELab
• Built agents that learned from user interactions to generate increasingly complex and personalized VR content in Unity

AI Intern | Deloitte, Singapore May 2023 – Aug 2023
• Built a retrieval-augmented generation (RAG)-based personalized assistant with LangChain, automating >50% of compliance tasks
• Presented to 300+ professionals across Southeast Asia; one of four interns firm-wide to receive the Deloitte Intern Impact Award

PROJECTS

[School Project] Commercial Space Analysis Platform
• Built AWS ETL pipelines consolidating data from 7+ sources into a database of 220+ Singapore malls, facilitating both Data-as-a-Service (DaaS) and clustering analysis for mall analytics and rental recommendations
• Selected as ‘Best Solution’ by PropertyGuru, Asia’s largest online property portal, for the project’s innovation and practical utility

[School Project] Choral music to sheet music converter
• Developed a novel music source-separation model, improving average signal-to-distortion ratio (SDR) from 2.88 to 2.92
• Co-authored a paper accepted at RAFDA 2024 and published in Springer’s Lecture Notes in AI Series

[Hackathon] Social Media Chatbot for Sales
• Built a generative-AI personalized assistant (FastAPI, LangChain, Redis, ChromaDB) with multi-turn conversational workflows, enabling end-to-end customer journeys from marketing to flight & ancillary purchases
• Achieved 1st runner-up among 287 teams in Singapore Airlines’ two-month global hackathon

SKILLS

• **Languages:** Python, SQL, C, C++, C#, Java, R, JavaScript, Go
• **AI/ML Tools:** NumPy, Pandas, PyTorch, TensorFlow, JAX, HuggingFace, scikit-learn, LangChain, vLLM, PySpark, spaCy, Tableau, Scrapy, Hadoop, Docker, Streamlit, Linux