# Zexi (Jesse) Zhang

📞 13264500190 | ✉ z13264500190@gmail.com | ⌂ github.com/nagi-ovo | 🌐 nagi.fun | 🔗 Zexi Zhang

## Education

**Beijing University of Technology**                                    2021/09 - 2025/06 (Expected)
Artificial Intelligence(senior at present), GPA 3.56/4.0                                    Beijing, China

- Main Courses: Machine Learning, Deep Learning, Natural Language Processing, Big Data Processing Technology
- Teaching Assistant for Python Programming (Fall 2024) & Natural Language Processing (Spring 2025)
- **TOELF: 94 | CET6: 588 | IELTS: 7.0**

## Selected Awards

- **Kaggle Silver Medal (Top 5% teams)**                                    2024/02 - 2024/04
  Team Leader                                    Kaggle Platform
  Led team in **LLM Prompt Recovery Challenge**, showcasing NLP expertise and leadership."
- **Provincial First Prize**, 2023 Contemporary Undergraduate Mathematical Contest in Modeling (CUMCM) | Modeling & Programming
- **Honorable Mention**, 2024 Mathematical Contest In Modeling (MCM), COMAP | Modeling & Programming

## Research Experience

**Beijing Key Laboratory of Multimedia and Intelligent Software Technology**                 2023/02 - 2023/09
Research Assistant                                    Beijing University of Technology, China

- Contributed to the development of MemoMusic 3.0 in AI4Health Team, an innovative system that enhances personalized music recommendation by considering music listening context and improves music generation by introducing music theory.
- Focused on improving the Transformer-based music generation framework, which was trained into three models for Classic, Pop, and Yanni music respectively.Implemented a novel approach using the dominant melody of music with expected Valence and Arousal values as a sample sequence to the model, with output adjusted according to music theory.
- Collaborated with an international research team, including members from University of Regensburg, Pandora, Cornell University, and University of California, Irvine.
- Research outcomes were published and presented at the 2023 IEEE International Conference on Multimedia and Expo Workshops (ICMEW), demonstrating improvements in enhancing listeners' emotional states and achieving better user satisfaction.

**Xiaoyan Li's Lab**                                    2024/09 - Present
Research Assistant                                    Beijing University of Technology, China
Core member, focusing on Incremental Learning with Vision Transformers (ViT).

## Internship Experience

**Pony.ai**                                    2024/04 - 2024/08
Software Development Intern                                    Beijing, China

- Served as an LLM & AI consultant for an automated annotation system, handling complex business scenarios and unstable data quality
- Conducted comprehensive research on LLM performance and deployment strategies, establishing evaluation metrics for diverse annotation tasks
- Applied advanced prompt engineering techniques and optimized workflows, achieving over 90% annotation accuracy for challenging, variable-quality data
- Implemented robust Python backend for data preprocessing and post-processing, reducing manual annotation workload by 80%
- Collaborated with cross-functional teams to integrate AI solutions into existing business processes, enhancing overall system efficiency

## Publication

**MemoMusic 3.0: Considering Context at Music Recommendation and Combining Music Theory at Music Generation**. L. Mou, Y. Sun, Y. Tian, Y. Sun, Y. Liu, **Z. Zhang**, R. He, J. Li, J. Li, Z. Li, F. Gao, Y. Shi and R. Jain. (2023). IEEE International Conference on Multimedia and Expo Workshops (**ICMEW 2023**). [Paper]

**BJUT-SWIFT Student Association**                                                      2023/07 - Present
Founder and maintainer

- **Founder and Team Leader** of  bjut-swift: A pioneering student-led organization dedicated to fostering a culture of knowledge sharing, innovation, and future technology (S.W.I.F.T.) at Beijing University of Technology.

- Spearheaded the development of an **open learning resource repository**  BJUT-Helper: Leveraged **Python** scripts and **GitHub workflows** to implement **CI/CD**, ensuring seamless deployment via mkdocs and GitHub Pages. This initiative exemplifies our commitment to democratizing access to educational resources.

- Orchestrated a strategic partnership between **bjut-swift** and the prominent **xiaoyuancat** online student community, significantly expanding our reach and impact. Led the creation and curation of diverse learning materials, including **LaTeX thesis templates, professional PPT designs, and innovative campus network tools**, demonstrating a keen understanding of student needs and technological trends.

- **Key Achievements**:
  ‣ Cultivated a vibrant community of like-minded individuals passionate about technology and education.
  ‣ Honed advanced skills in **team leadership, agile project management, and collaborative development**.
  ‣ Fostered an environment of continuous learning and innovation, directly benefiting hundreds of students.
  ‣ Demonstrated exceptional initiative and vision in identifying and addressing gaps in student resources.

## Technical Skills

- **Program Languages:** Python, C/C++, JavaScript/TypeScript
- **Frameworks and Tools:** Pytorch, OpenAI Triton, Docker, Next.js
- **AI framwokrs:** Natural language Processing (Llama, GPT-2) | Reinforce Learning(RLHF) |

## Open-Source Contributions

**Crossbell-Box/xLog** &#9711;                                                                Contributor
- Using **Typescript** to add new features and fix bugs for this blog platform that I mainly use,
- Actively participate in this web3 ecosystem, publish multiple high-quality blog posts  nagi.fun, mainly focusing on **deep learning, reinforcement learning, language models(LLMs) and programing**

**Datawhalechina/LLM-Deploy** &#9711;                                                          Contributor
- Contributed to open-source LLM deployment learning materials. Focused on concurrency improvement using vLLM. Specialized in network communication, decoding, and distributed inference optimization
- Provided practical model and service optimization strategies

**Personal Projects**                                                            ★ 461 &#9711; Nagi-ovo
- &#9711; **CHSI-Converter** (★277): Automated tool that instantly converts Chinese academic credentials into English for education certification. Built with Flask and containerized using Docker for easy deployment and scalability.
- &#9711; **CRAG-Ollama-Chat** (★66): Implemented CRAG (Corrective RAG) using Langgraph, significantly reducing LLM inaccuracies and hallucinations. Integrated external knowledge via web search tools and employed lightweight retrieval evaluator for document relevance and reliability. Compatible with OpenAI API and local LLMs via Ollama.
- &#9711; **Cherno-CPP-Notes** (★105): Keynotes and insights from modern C++ courses, including personal interpretations and opinions. Received positive feedback on forums for in-depth understanding and practical tips.

## Volunteer Experience

**Winter Olympics City Volunteer Team**                                                         2022/01
Volunteer                                                                               Beijing, China
- Served as a dedicated volunteer near Shougang Park, a key venue for the 2022 Winter Olympics in Shijingshan District, Beijing.
- Enhanced visitor experience by providing essential services, including distributing hot beverages, offering informative brochures, and assisting with photo opportunities.

- Demonstrated strong interpersonal skills and cultural sensitivity while interacting with diverse international tourists, contributing to a positive image of the host city.