Kaiqiang Zheng . Jarvis Consulting

I am a Software and AI Engineer with a Master's degree in Computer Engineering from the University of Toronto and a Bachelor's degree in Automation Engineering from McMaster University, specializing in Applied AI, Software Development and Cloud Computing. As an AI Web Application Research Associate at UofT, I developed and maintained responsive web applications using HTML, CSS, JavaScript, and React to support AI research. I integrated Flask-based Python backends with real-time data visualization, built APIs to fetch and render AI model outputs, implemented multi-agent Large Language Model pipelines using Python, LangChain and OpenAI API, and collaborated with graduate researchers to design intuitive UI/UX and SQL-backed data storage. I am passionate about building robust, interactive software systems and applying AI-driven solutions to real-world problems.

Skills

Proficient: Python, AI/NLP (LLMs, Multi-agent Architecture, OpenAI API, Prompt Engineering, Langchain), Machine/ Deep Learning (TensorFlow, Pytorch), Linux/Bash, RDBMS/SQL, Java, HTML/CSS, Databricks, Git, Agile/Scrum, LaTeX

Competent: C++, JavaScript, Docker, Cloud computing (AWS, Azure, GCP), REST APIs, Flask, React, Angular, Django

Familiar: MATLAB, C, Hadoop, Postman, Node.js, Figma

Jarvis Projects

Project source code: https://github.com/jarviscanada/jarvis_data_eng_KaiqiangZheng

Cluster Monitor [GitHub]: Developed a Linux cluster monitoring system for Jarvis' LCA team using Docker-contained PostgreSQL to store node performance data. Implemented Bash scripts to automate container lifecycle, database table creation, and data insertion from Linux hardware metrics. Scheduled one-minute interval data collection via crontab and conducted manual testing on GCP virtual machines running Rocky Linux 9. Leveraged Git for version control, centralizing monitoring and reducing manual overhead, enabling administrators to efficiently track system performance and resource usage across multiple clusters.

Highlighted Projects

SherlockGPT: Root Cause Analysis with Large Language Model [GitHub]: Developed an automated Root Cause Analysis (RCA) engine leveraging large language models (GPT-4 via OpenAI API) and prompt engineering to identify incident root causes from cloud system logs and metrics. Designed a multi-agent architecture that iteratively analyzes and ranks potential causes. Utilized the LEMMA-RCA dataset for evaluation, enabling improved incident triage and decision-making efficiency.

Sentiment Analysis for Customer Service Ticket Analysis [GitHub]: Built and deployed a sentiment analysis application in Docker to classify customer service ticket sentiments. Integrated OpenAI¡¯s NLP API for text analysis and Gradio for an interactive web interface, enabling real-time visualization of sentiment trends and improving customer feedback insights.

Full Stack Semester Project: Designed and implemented a full-stack web application using Django, Python, and MySQL for backend logic and data management. Built a responsive frontend using HTML, CSS, Bootstrap, and JavaScript. Containerized the application with Docker for consistent deployment, demonstrating proficiency in full-stack development and cloud deployment workflows.

Professional Experiences

Software Developer, Jarvis (2025-present): Developed and maintained scalable software solutions within an Agile/Scrum framework using Linux, Bash, SQL, Git, Java, and Docker. Contributed to multiple projects, designing and implementing efficient backend systems, optimizing database performance, and streamlining deployment processes. Participated in sprint planning, daily stand-ups, and retrospectives to ensure iterative progress and high-quality delivery. Leveraged big data technologies such as Apache Spark and Databricks to handle large datasets and build robust data processing pipelines.

AI Web Application Research Associate, University of Toronto SG (2024): Developed and maintained responsive web applications using React, HTML, CSS, and integrated Python Flask backends to visualize AI model outputs in

real time. Collaborated with graduate researchers to design intuitive UI/UX interfaces and integrate SQL databases for experimental data storage. Implemented multi-agent NLP pipelines using Python and OpenAI API, supported scalable AI workflows, and followed Git-based version control and regular code reviews to ensure robust software practices.

Design Technician - Co-op, University of Toronto SG (2022-2023): Utilized software tools to code programs and perform calculations, creating accurate design drawings for engineers and ensuring compliance with specifications. Supported busbar design through programming to assist assembly processes. Managed administrative tasks including enrolling new employees and maintaining records, contributing to smooth operational workflows.

Education

University of Toronto (2024-2025), Master of Engineering, Electrical and Computer Engineering - GPA: 3.83/4.0

McMaster University (2019-2023), Bachelor of Engineering Technology, Automation Engineering - Graduated with Distinction - Scholarship - Dean's List (2021, 2022) - GPA: 3.70/4.0

Miscellaneous

- AWS Academy Cloud Foundations Badge (2023)
- FANUC CERT HandlingTool Operations and Programming (2022)
- 3rd Place for Best Design in the International Design and Engineering Education Association (IDEEA) Competition (2023)
- Fall ECE Graduate Orientation: Assisted the ECE Graduate Office by participating in a mix-and-mingle reception, addressing student inquiries, and providing information and guidance about the program.
- Talent Show: Completed a wonderful performance by playing the guitar while maintaining a hard study state.
- Teaching Assistant: Provided one-on-one tutoring in mathematics to summer students, covering algebra and calculus.