Kexuan Zhang

+1 437-361-2016 | zhangkexuan0514@outlook.com | Toronto, ON

TECHNICAL SKILLS

- Programming Language: **Java, JavaScript, Python, C++, C, Ruby, .NET, TypeScript, SQL, Node.js, HTML/CSS**;
- AI/ML Frameworks: **TensorFlow, PyTorch, NumPy, OpenAI, HuggingFace, Graph RAG, Scikit-Learn**;
- Software Development Tools: **Django, Flask, React, Azure, AWS, GCP, IBM Cloud, NoSQL, Git, Kubernetes, Docker**;

EDUCATION

University of Toronto Toronto, ON

Master of Electrical and Computer Engineering, Emphasis in Analytics Sept. 2024 – Apr. 2025(Expected)

- CGPA: 3.94/4.0.
- Relevant coursework: *Application of Natural Language Processing, Empirical Software Engineering, Convex Optimization, Cloud Computing, Wearable AI*

University of Toronto Toronto, ON

Bachelor of Science – Computer Science Specialists, Focus in Artificial Intelligence Sept. 2020 – Jun. 2024

- CGPA: 3.75/4.0. Honour: New College Council In-Class Scholarship 2022-2023.
- Relevant coursework: *Deep Learning and Neural Network, Natural Language Computing, Visual Computing, Advanced Algorithm, Data Structure, Database, Operating System,*

WORK EXPERIENCE

Machine Learning Engineer Remote

EUAI Education Jun. 2024 – Dec. 2024

- Developed an AI-powered teaching platform for partner universities in East Asia, integrating AI modules to support professors in teaching and enhance students' learning in finance and accounting.
- Led the design and implementation of multiple chat generation pipelines using the **GraphRAG** technique, integrating the company's knowledge base with **Azure** cloud services, including **AI Search Service**, **Blob Storage**, and **AI Studio**.
- Developed **RESTful APIs** for the AI modules using **Flask**, successfully hosting the solution on an **AWS cloud server** for production use. Due to the independence and versatility of this functionality, the company decided to encapsulate the APIs as a standalone **PaaS** (Platform-as-a-Service) product, enabling broader application and potential for additional revenue streams.

Backend Developer Toronto, ON

Bellwoods Strategic Capital Jan. 2023 – Apr. 2023

- Maintained and extended a **Django**-based mental health application Thrive, for users to enroll in "journeys" which include a series of "quests" to complete. Completing the "quests" designed by company psychologist team helps users to improve mental health and enhance overall happiness.
- Designed and implemented a robust quest/journey **matching algorithm** using collaborative filtering and content-based filtering, improving matching accuracy by 67%.
- Developed and optimized the progress history feature, utilizing database management systems (DBMS) such as **PostgreSQL** and cloud datastore **Heroku** to store and retrieve user data efficiently, and deployed the application using **Docker**.

PROJECT & RESEARCH EXPERIENCE

Research Assistant – University of Toronto

Need Satisfier Extraction Network Sept. 2023 – Apr. 2024

- Participated in the research project to develop a novel **neural network**, SNM-I, which aimed to extract "need satisfiers" from service or program descriptions provided by various agencies and institutions.
- Researched and formulated a parallel structured NLP pipeline, comprising an extractive pipeline to identify potential need satisfiers and a generative pipeline to understand the sematic meaning and uncover hidden need satisfiers.
- Formulated procedural guidelines and **ontology-based reasoning** for the **generative model**, instructing the pipeline to identify characteristic changes introduced by the service.
- Conducted **fine-tuning** on **pre-trained large language models**, including **RoBERTa-large, Liama 2, and GPT-3.5 turbo**, to optimize performance. Executed comprehensive efficiency and accuracy analyses to assess model effectiveness.

Software Developer

Utinder Student Networking Application Sept. 2022 – Dec. 2022

- Designed and developed a comprehensive social network application called Utinder for students, incorporating essential features such as a social forum and events calendar, enabling users to create posts, comment on posts, and engage in discussions.
- Implemented the backend in **Java** and connected the application to **Firebase Cloud Datastore** for secure data storage. Developed user authentication and authorization mechanisms to safeguard user data and maintain privacy.
- Built the app's frontend using **Android Studio**, creating a user-friendly interface for seamless interaction.