

YUJIA ZHANG, PH.D.

Edmonton, Canada ◊ yujia10@ualberta.ca ◊ [Linkedin: Yujia-Zhang-PhD](#) ◊ [Github: Yujia](#)

SUMMARY

Machine learning researcher with 8+ years of experience in developing and deploying ML models, specializing in knowledge graphs and language models. Skilled with 6+ in leveraging deep learning frameworks (PyTorch, TensorFlow) and cloud platforms (GCP) to drive promising impact.

EDUCATION

Ph.D. of Computer Engineering , GPA 4/4, University of Alberta	Sept. 2019 - Nov. 2024
Master of Computer Engineering , GPA 4/4, Chonbuk National University	Aug. 2016 - Aug. 2018
Bachelor of Engineering , GPA 3.7/4, Shanghai University of Engineering Science	Sept. 2011 - July 2015

EXPERIENCE

Machine Learning Resident <i>Amii</i>	March 2025 - <i>Edmonton, CA</i>
---	-------------------------------------

- Delivered **end-to-end LLM** projects, from initial client scoping and data analysis to building scalable models that provide actionable insights.
- Developed a **Retrieval-Augmented Generation (RAG)** system and exposed it via a Flask REST API. Containerized the application using **Docker** to ensure seamless deployment and environment consistency.
- Drove AI innovation by translating cutting-edge research in domains like **Agentic LLMs** into practical, industry-ready applications and solutions.

Data Analyst <i>Nanoprecise Sci Corp</i>	May 2019 - Oct 2019 <i>Edmonton, CA</i>
--	--

- Implemented **exploratory data analysis (EDA)** to analyze sensor signal data, enabling predictive maintenance for sensors with a **95% accuracy rate** in failure detection by **Random Forest** Model.
- Migrated and optimized firmware code to Python, improving the efficiency and scalability of the data processing pipeline.
- Conducted **comprehensive sensor data analysis** to identify patterns and anomalies, providing actionable insights to prevent equipment failures.
- Created dynamic **visualization dashboards** to communicate complex analytical results effectively, presenting findings to both technical and non-technical stakeholders, including customers.

R&D Product Engineer YUNGTAY ELEVATOR EQUIPMENT (CHINA) CO., LTD.	May 2015 - July 2016 <i>Shanghai, China</i>
---	--

- Led the **data collection and quality assurance process** for elevator measurement data, ensuring compliance with strict industry standards and performance benchmarks.
- Conducted **statistical analysis** and implemented quality control techniques to verify that product measurements met design and regulatory requirements.
- Authored 26 detailed **technical reports** summarizing analysis results, providing actionable recommendations to improve product quality and reliability.
- Collaborated with cross-functional teams, including design and manufacturing, to address discrepancies and optimize product performance based on data insights.

PROJECTS

Knowledge Graphs Construction via Large Language Models (LLM). <i>University of Alberta</i>	June 2023 - May 2024 <i>Edmonton, CA</i>
---	---

- Problem:** To develop an efficient system for constructing KGs from unstructured text by fine-tuning LLMs.
- Team Collaboration:** Collaborated closely with team members to align project goals, share insights, and set a timeline to follow.
- Model Deployment:** Designed and implemented a scalable LLM deployment pipeline using **Hugging-Face Transformers**. Optimized model inference for **Llama3** and **Mistral** through Autotrain, ensuring a professional setup.
- Scalable Training:** Developed scalable fine-tuning workflows for LLM, achieving a **30% reduction in training time** while handling diverse datasets such as WebNLG and SKE.
- Key Findings:** **smaller LLMs of 7b parameters can achieve or exceed the performance of big LLMs like GPT-4** in specific tasks such as triple extraction.

Hierarchical Analysis for Knowledge Graphs (KGs).

University of Alberta

May 2020 - Dec. 2023

Edmonton, CA

- Problem:** To develop a generative model for knowledge graphs.
- Model Development:** Developed a **non-parametric hierarchical generative model** for KGs, utilizing probabilistic topic modeling techniques to achieve **structured clustering** without preset assumptions.
- Inference Optimization:** Implemented an efficient **Gibbs sampling** method for scalable inference in Python, enhancing model performance and processing speed.
- Evaluation:** Conducted **quantitative and qualitative evaluations** on standard datasets, demonstrating comparable performance to established hierarchical clustering methods.
- Publications:** Two papers based on this work have been published.

Object Detection and Tracking.

Chonbuk National University

Aug. 2016 - Aug. 2018

Jeonju, South Korea

- Problem Solving:** Identified key challenges in existing models and developed innovative solutions, resulting in significant improvements in tracking accuracy and robustness.
- Algorithm Research:** Investigated advanced object tracking and recognition techniques using **Siamese networks, LSTM, CNN, and RNN** algorithms implemented in Python with TensorFlow.
- Performance Enhancement:** Optimized model architectures and training procedures, achieving an increase in **mean Average Precision (mAP)** on standard benchmark datasets.
- Publication:** Authored multiple research papers detailing novel approaches and empirical results, published in peer-reviewed conferences and journals.

TECHNICAL SKILLS

Database: SQL, MySQL, NoSQL, MongoDB, Neo4j

Deep Learning Frameworks: Tensorflow, Pytorch, Keras, Scikit-learn

Development Tools: Git, Docker, Jupyter

ML Operations: Model Deployment, Inference Optimization

Programming Languages: Python, Matlab, Java, C++

Visualization: Matplotlib, Seaborn, Tableau

Cloud Services: AWS, Azure, Google Cloud Platform/ AI Studio

EXTRA-CURRICULAR ACTIVITIES

- Volunteer for the energy Adventure in TELUS World of Science May 2024
- TA in ECE 493 Software Systems Design Project Jan. 2024 - Apr. 2024
- TA in ECE 627 Intelligent Web Jan. 2023 - Apr. 2023

- TA in ECE 447 Data Analysis and Machine Learning for Engineers Jan. 2023 - Apr. 2023
- TA in ENCMP 100 Computer Programming for Engineers Jan. 2021 - Apr. 2021
- ICTAI 2022 Conference Publicity Chair Oct 2022

PUBLICATIONS

- **PhD Thesis Paper:** Constructing Knowledge Graphs with Language Models and Learning Hierarchies from Graphs using Probabilistic Topic Modeling, Sept. 2024
- **Journal:** Construction of Topic Hierarchy with Subtree Representation for Knowledge Graphs, *Axioms* 2025, 14, 300. <https://doi.org/10.3390/axioms14040300> Apr. 2025
- Fine-tuning Language Models for Triple Extraction with Data Augmentation, *Knowledge Graphs and Large Language Models (KaLLM) Workshop*, 62nd Annual Meeting of the Association for Computational Linguistics (ACL 2024) Aug. 2024
- Hierarchical Topic Models for Knowledge Graphs, *19th European Semantic Web Conference (ESWC 2022), Hersonissos, Greece* June 2022
- Robust Object Tracking with Convolutional Gated Recurrent Networks, *International Symposium on Information Technology Convergence (ISITC 2017)*, Shijiazhuang, China May 2017
- Deep Residual Convolutional Network for Natural Image Denoising and Brightness Enhancement, *2018 International Conference on Platform Technology and Service (PlatCon-18)*, Jeju, Korea May 2018
- Object Tracking via Spatial-Temporal Learning Combined Correlation Filter, *Smart Media 2018*, Seoul, Korea Feb. 2018
- **Master Thesis Paper:** Efficient Correlation Matching-Based Real-Time Object Tracking with Spatiotemporal Information, June 2018

AWARDS

- J Gordin Kaplan Graduate Student Award, University of Alberta Aug 2024
- Brain Korea 21 Scholarship Aug 2017
- BECKHOFF Scholarship by SUES, Beckhoff Automation Co., Ltd. Bi Fu automation equipment trading (Shanghai) Co., Ltd. July 2014