Zerui Wang

+1 438 408 6668 | wangzerui418@gmail.com | Portfolio: xai-hub.com | GitHub | LinkedIn | Montréal, Canada

Education

Concordia University May 2020 – Present
Ph.D. in Computer Engineering *Canada*
- Research Focus: AI/Machine Learning, Cloud AI Service, Transformer, NLP, Computer Vision, Video Model, Explainable AI
- Course Lecturer and Teaching Assistant: COEN6311 Software Engineering, COEN6313 Programming on Cloud, COEN244 Programming Methodology, COEN6731 Distributed Software Systems

Technical University Dortmund Oct 2014 – Jan 2018
M.Sc. in Process System Engineering *Germany*
- Specialization: System Modelling, Dynamic Simulation, System Control Programming, Parallel Computing

China University of Mining and Technology Sep 2010 – Jul 2014
B.Sc. in Process System Engineering *China*
- Studies: Advanced Mathematics, Linear Algebra, Physics, Statistics, Process Design, Programming

Professional Experience

Intellipro Group Inc Apr 2023 – Dec 2023
AI Engineer & Trainee Supervisor (Intern) *Toronto, Canada*
- Developed a chatbot using open-source LLMs and APIs, integrating technologies such as **prompting engineering, retrieval-augmented generation (RAG), tokenization, and vector databases** to enhance the answer with domain knowledge.
- Participated in the design and development of a **recommendation system** for candidates. Specifically, responsible for fine-tuning the LLAMA model to extract key feature keywords from candidate profiles, enabling accurate assessment of job position suitability.
- Adopted and fine-tuned LLMs for in-house tool applications, enhancing model alignment and performance for specialized HR office tasks. Designed and implemented innovative **office automation tools**, including context-aware email draft generation, semantic search for regulations, and automated meeting summary generation.
- Mentored and supervised a team of programming developer trainees, providing guidance and support for their internship.

Concordia University May 2020 – Present
Doctoral Researcher *Montréal, Canada*
- Developed an advanced **Explainable AI (XAI) framework** to quantify feature contribution explanations for models handling diverse data types (tabular, text, image) and cloud AI systems, including search ranking, code vulnerability detection, and computer vision cases. Published on top-tier conference ICSE.
- Engineered an innovative **microservice-based, open-API architecture** to integrate XAI process operations into black-box cloud AI services, enabling explainability without exposing underlying model structures.
- **Enhanced cloud AI service** performance and explainability for Microsoft Azure AI, Google Cloud Vertex AI, and Amazon Web Services. Experiments published in top-tier journal IEEE Transaction on Cloud Computing.
- Designed and implemented an **assessment framework** for **AI service quality assurance**, evaluating AI models under adversarial conditions to ensure robustness and explanation accuracy in real-world operational scenarios. Published on the IEEE SSE.
- Developed and deployed an AI-driven **financial data analysis tool**, enabling stock price forecasting and comprehensive analysis reports. The project is based on the financial data APIs to retrieve real-time market information and historical financial data.

Paper Publications

First Author Publications:

- **Wang, Z.**, Liu, Y., Thiruselvi, A.A., Hamou-Lhadj, A. (2024). "XAIport: A Service Framework for the Early Adoption of XAI in AI Model Development." ***International Conference on Software Engineering (ICSE)***. DOI:10.1145/3639476.3639759
- **Wang, Z.**, Liu, Y., Huang, J. (2024). "An Open API Architecture to Discover the Trustworthy Explanation of Cloud AI Services." ***IEEE Transactions on Cloud Computing***. DOI:10.1109/TCC.2024.3398609
- **Wang, Z.**, Liu, Y. (2024). "Cloud-based XAI Services for Assessing Open Repository Models Under Adversarial Attacks." ***IEEE International Conference on Software Services Engineering***. arXiv:2401.12261
- **Wang, Z.***, Huang, J.*, Li, D., Liu, Y. (2022). "The Analysis and Development of an XAI Process on Feature Contribution Explanation." ***IEEE International Conference on Big Data***. 5039-5048. DOI: 10.1109/BigData55660.2022.10020313

Co-authored Publications:

- Neghawi, E., **Wang, Z.**, Huang, J., Liu, Y. (2023). "Linking Team-level and Organization-level Governance in Machine Learning Operations through Explainable AI and Responsible AI Connector." ***IEEE 47th Annual Computers, Software, and Applications Conference (COMPSAC)***. 1223-1230. DOI: 10.1109/COMPSAC57700.2023.00185
- Li, D., Liu, Y., Huang, J., **Wang, Z.** (2023). "A Trustworthy View on Explainable Artificial Intelligence Method Evaluation." ***IEEE Computer***, 56(4), 50-60. DOI: 10.1109/MC.2022.3233806

Presentations

- **Tutorial Workshop Lead:** "Develop Explainable AI Services on Cloud Computing and Open Source Models" at ***34th International Conference on Collaborative Advances in Software and Computing (CASCON)***. (2024)
- **Poster Presenter:** "Cloud-based XAI Services for Assessing AI Models Under Adversarial Attacks" at ***Software Engineering for Machine Learning Applications (SEMLA) International Symposium***. (2024)
- **Tutorial Presenter:** "How to Design and Launch a Software-as-a-Service (SaaS) on Cloud Computing" at ***Concordia University 4th Annual Digital Skill-Share Days Event***. (2024)
- **Research Presenter:** "Design Explanation Microservices and Provenance: A Case Study of Explaining Cloud AI Service" at ***The 36th Canadian Conference on Artificial Intelligence***. (2023)

Technical Skills

Programming & Software Development: Python; C/C++; Algorithm Design; Version Control (Git); CI/CD; RESTful APIs; Swagger; Microservices Architecture

Data Processing & Analysis: SQL and NoSQL Databases; Data Preprocessing And Feature Engineering; Data Augmentation; Data Visualization (Matplotlib; Seaborn)

Machine Learning & Deep Learning: PyTorch; TensorFlow; MXNet; CNNs; Transformers; Hyperparameter Tuning;

Natural Language Processing (NLP): Word Embeddings; LLMs; Sentiment Analysis; Attention Mechanisms; PEFT

Computer Vision: Image/Video Classification; Vision Transformer; Supervised CNNs/ViTs Training; Saliency Map Extraction; Semantic Segmentation; Multi-modal Model CLIP; Adversarial Attack, Robustness Evaluation

Explainable AI (XAI): SHAP; LIME; Feature Importance Analysis; Model-agnostic Explanation Methods; Fidelity Evaluation; Fairness And Bias Detection

Cloud Computing & DevOps: AWS (EC2; S3; SageMaker); GCP (Compute Engine; BigQuery; Vertex AI); Azure (VMs; Blob Storage; Azure ML; OpenAI APIs); Docker; Kubernetes; Infrastructure As Code (Terraform)

MLOps & Experiment Management: Apache Airflow; Kubeflow; Weights & Biases; Model Versioning; Experiment Tracking; Model Monitoring; HuggingFace

AI Ethics & Responsible AI: Fairness in AI; Bias Mitigation; Ethical AI Frameworks; AI Governance

AI Research & Development: Implementing SOTA Techniques; Adapting Research Papers To Practical Applications; Experimental Design; Experiment Reproduction; Technical Writing And Documentation

Awards, Certifications and Professional Memberships

- **Project Management Professional (PMP)®** Certification, Project Management Institute
- PhD Entrance **Scholarship**, Concordia University
- Student Member, Association for Computing Machinery (ACM)
- Student Member, IEEE Computer Society