

Mohammad Arvan

Chicago, IL mo-arvan.github.io linkedin.com/in/mo-arvan

Education

University of Illinois at Chicago (UIC)

Aug 2024

Ph.D., Computer Science

GPA: 4.0/4.0

- **Dissertation:** Machine Learning and Open Science: On Risks and Challenges

Qazvin Islamic Azad University (QIAU)

July 2016

B.Sc., Software Engineering

Ranked 2nd in class of 180 students

Skills

- **Libraries and Technologies:** PyTorch, NumPy, Pandas, Huggingface Transformers, PyTorch Lightning, Numba, TensorFlow, Keras, JAX, scikit-learn, SciPy, Streamlit, Matplotlib, Spacy, NLTK, Docker, Kubernetes, AWS, GCP, Azure, Linux/Unix, ssh, Git

- **Languages:** Python, C++, SQL, Octave/MATLAB, C#, Java, R

Experience

Research Assistant

Aug 2018 to Aug 2024

UIC Natural Language Processing Laboratory

Chicago, IL

- Pioneered research in applying Large Language Models (LLMs) to unstructured clinical notes at the University of Illinois Hospital. Achieved up to 90% F1 score in detecting houseless populations, cardio-oncology issues, cervical cancer screening, and COVID-19 symptoms, significantly improving patient identification and care strategies.

- Designed and demonstrated a proof-of-concept for distributed training of neural networks using heterogeneous compute units over unreliable connections. This paradigm enables training of extremely large networks via volunteer computing, potentially democratizing access to high-performance AI training resources.

- Implemented, trained, and evaluated LSTM and transformer models for specific tasks such as language modeling, machine translation, and mathematical question answering. These models achieved state-of-the-art performance, contributing to advancements in sequence processing methodologies.

- Achieved 80% accuracy in retrieval-augmented question answering on the SQuAD dataset by implementing transfer learning with BERT and utilizing corpus-level information retrieval.

Skills: Foundation Models, Generative AI, Distributed Training, Transformer, Linear Algebra, Calculus, Probability, Deep Learning, Neural Networks, Fine-tuning, AI/ML Frameworks, RLHF.

Software Engineer

Aug 2016 to Aug 2018

Kara Intelligent System

Iran

- Designed and implemented interactive Tableau dashboards with advanced filtering and drill-down capabilities for monitoring network performance, leading to a 10% reduction in response time to system failures. These dashboards provided real-time insights into traffic patterns and system anomalies across a telecom network supporting over 40 million users.

- Achieved an 85% reduction in data processing time for ETL (Extract, Transform, Load) workflows, decreasing the processing time from 168 hours to less than 24, by utilizing parallel processing and data streaming techniques.

- Collaborated with finance, marketing, and operations teams to identify critical business and performance requirements, translating them into detailed technical specifications.

Skills: Data Visualization, Data Cleaning, API Development, Cloud Computing, Statistical Analysis.

Co-Founder and Developer

Aug 2015 to Aug 2016

Indooria Startup

Iran

- Led the design and development of a real-time indoor positioning and mapping system for airports, shopping malls, and warehouses. Utilized advanced triangulation techniques, wireless signal modeling, and robust filtering methods to achieve precise indoor location tracking with an accuracy of 5 meters.

- Utilized optimization and batch rendering to improve the map rendering framerate by 100%, increasing it from 30 FPS to over 60 FPS, enhancing user experience.

- Employed tailored communication strategies such as simplified visual aids and detailed technical documents, for a diverse audience including investors and engineers to ensure clear understanding and alignment, leading to successful project implementation and stakeholder satisfaction.

Skills: Data Structures, Software Architecture, Object-Oriented Programming, Technical Leadership, Agile Development

Research Assistant

Aug 2013 to Aug 2015

Mechatronic Research Laboratory

Iran

- Successfully implemented AI algorithms, including path planning, motion planning, and behavior control, in robotics, leading to winning first place in the 2015 international RoboCup Rescue competition.

- Designed and developed feature engineering processes and evaluated machine learning models, including AdaBoost, SVM, Decision Trees, and Random Forests, for real-time edge device object detection achieving a 90% accuracy rate.

Skills: Machine Learning, Computer Vision, Data Analysis, Problem Solving, Rapid Growth and Learning, C++, Artificial Intelligence

Publications

- **Arvan** et al. "ReproHum #0712-01: Human Evaluation Reproduction Report for "Hierarchical Sketch Induction for Paraphrase Generation" **HumEval 2024**

- **Arvan** et al. "Human Evaluation Reproduction Report for Data-to-text Generation with Macro Planning." **HumEval 2023**

- **Arvan** et al. "Investigating Reproducibility at Interspeech Conferences: A Longitudinal and Comparative Perspective." **INTERSPEECH 2023**

- Popović et al. "Exploring Variation of Results from Different Experimental Conditions." **ACL Findings 2023**

- **Arvan** et al. "Linguistic Cognitive Load Analysis on Dialogues with an Intelligent Virtual Assistant." **CogSci 2023**

- Anya Belz et al. "Missing Information, Unresponsive Authors, Experimental Flaws: The Impossibility of Assessing the Reproducibility of Previous Human Evaluations in NLP." **Insights 2023**

- **Arvan** et al. "Reproducibility in Computational Linguistics: Is Source Code Enough?" **EMNLP 2022**

- Haghighat et al. Effects of an Intelligent Virtual Assistant on Office Task Performance and Workload in a Noisy Environment." **Applied Ergonomics**, 109, 103969.

- Haghighat et al. "Human Interaction with Intelligent Virtual Assistant in a Noisy Environment." **HFES 66th**

- **Arvan** et al. "Reproducibility of Exploring Neural Text Simplification Models: A Review." **INLG 2022**

Honors and Awards

- Finalist in the 2024 **Three Minute Thesis (3MT)** Competition at the University of Illinois at Chicago.

- Recipient of the 2020 **Provost's Graduate Research Award**. University of Illinois at Chicago (\$5000)

- Innovative User Interface Award, **RoboCup World Championship**, Rescue Robot League. Hefei, China, 2015

- Ranked 1st, **RoboCup World Championship**, Rescue Robot League. Hefei, China, 2015

Certificate

- "Fundamentals of Accelerated Computing with CUDA C/C++," Issued by **NVIDIA Deep Learning Institute**, 2023