Amirreza Sokhankhosh

431-293-6515 amirreza.skhn@gmail.com LinkedIn GitHub Portfolio

Highlight of Qualifications

- **Generative AI & LLMs:** Developed advanced AI systems leveraging **GPT-4** and other **large language models** to automate complex tasks, achieving up to a **30% reduction** in processing time and improving efficiency in project outputs.
- **Machine Learning Engineering:** Extensive experience in **developing** and **deploying machine learning models** on **Google Cloud Platform** (GCP), employing tools like **Vertex AI** and **BigQuery** to ensure high-quality model performance in production environments.
- **Deep Learning & Neural Networks:** Proficient in building and optimizing **deep learning architectures** with frameworks such as **TensorFlow** and **PyTorch**, enhancing model reliability and accuracy through thorough evaluation and optimization practices.
- **NLP & Computer Vision:** Designed intelligent systems using **natural language processing** and **computer vision** techniques, including a **paper summarizer** that improved summarization efficiency by **30%**, showcasing expertise in handling complex text-based datasets.
- **Cloud & DevOps:** Skilled in employing **MLOps** practices, integrating CI/CD pipelines for seamless deployment and maintenance of machine learning applications, addressing scalability and operational reliability across distributed systems.
- **Collaboration & Leadership:** Demonstrated ability to lead cross-functional teams and define project objectives, successfully completing high-impact projects while effectively communicating complex technical concepts to diverse stakeholders.

Experience

University of Manitoba September 2023 – July 2025
Graduate Research Assistant *Winnipeg, Manitoba, Canada*

- Designed and developed **novel distributed AI architectures** addressing challenges in distributed learning, utilizing **TensorFlow** and **PyTorch** for model development.
- Implemented and optimized a blockchain consensus mechanism **(PoCL)** to enhance resource efficiency and security in federated learning, reducing communication overhead by **85.2%**.
- Authored and contributed to research published in top-tier IEEE venues, showcasing expertise in **problem-solving** and **technical communication** to diverse audiences.
- Conducted thorough model evaluations and optimizations to ensure high-quality outputs, aligning with best practices in **MLOps** for deploying models in production environments.
- Collaborated with cross-functional teams to define project goals and build scalable machine learning architectures for complex text-based datasets.

K. N. Toosi University of Technology June 2021 – August 2022
Research Assistant

- Led a research team in data cleaning and preprocessing for over **10GB** of raw data, ensuring high-quality datasets necessary for advanced machine learning analyses.
- Utilized **R** and advanced causal inference techniques to analyze data, producing insights that highlighted significant biases, aligning well with the evaluation aspects of machine learning.
- Facilitated cross-team collaboration to tackle complex projects, honing strong **problem-solving** skills necessary for the role of a machine learning engineer.
- Presented analytical findings to stakeholders, demonstrating the ability to convey complex technical concepts effectively to both technical and non-technical audiences.
- Supported academic initiatives through the design of educational material that integrated theoretical knowledge with practical applications in machine learning.

Bobo May 2024 – August 2024
Full-stack Developer Intern *Winnipeg, Manitoba, Canada*

- Accelerated product development by designing **RESTful APIs** to enhance backend functionality, ensuring data management aligned with user interface requirements.
- Automated data integration using a Python script, enabling efficient population of databases, demonstrating strong **Python** proficiency and an eye for optimization.
- Collaborated with cross-functional teams in an **agile environment**, utilizing tools like **Jira** for task management and improving workflow efficiency.
- Maintained focus on delivering scalable solutions, ensuring that the integration of APIs met the needs of various front-end components.
- Engaged in hands-on learning and implementation of cloud-based technologies, broadening familiarity with machine learning deployment practices on platforms such as GCP.

Projects

Paper Summarizer | *Python, PyTorch, Flask, Detectron2, LLaVA* | Code

- Developed an intelligent academic paper summarization system using **natural language processing** and **computer vision** techniques that improved summary generation efficiency by 30%.
- Implemented **deep learning** algorithms with **PyTorch** to extract and analyze document elements, leveraging **large language models** for comprehensive text summarization.
- Engineered a scalable **Flask API** that supported RESTful communication, enabling multi-modal analysis of academic papers and reducing processing time by over 50%.

CIFAR-10 Generative Model Evaluation | *Python, PyTorch, ResNet-50, Custom Metrics* | Code

- Developed a robust evaluation framework using **PyTorch** and **ResNet-50** to assess generative models on the CIFAR-10 dataset.
- Implemented **custom metrics** (Precision, Recall, Generalization Rate) to ensure **high-quality model evaluation** using nearest-neighbor analysis.
- Achieved a notable increase in evaluation reliability with metrics demonstrating **over 90% accuracy** in matching real and generated samples.

MarkMate | *Django, Flask, GPT-4, PostgreSQL, Python* | Code

- Developed an **AI-powered grading** system utilizing **GPT-4** for **automated assignment evaluation**, enhancing accuracy and efficiency for instructors.
- Implemented a **microservices architecture** using **Django** and **Flask**, enabling scalable and maintainable services that supported large-scale usage by multiple users.
- Achieved a **30% reduction in grading time** for instructors, allowing for more effective use of educational resources and improved student feedback quality.

PyFed | *Python, TensorFlow, scikit-learn, BigQuery, TensorBoard* | Code

- Developed a lightweight federated learning framework with a **client-server architecture** using **Python** and **TensorFlow**, enabling efficient model training across distributed systems.
- Implemented the **FedAvg policy** using **multi-threaded** processes, optimizing communication and enhancing the scalability of federated learning tasks.
- Achieved a **30% reduction in training time** for federated learning tasks while maintaining model accuracy, demonstrating the framework's effectiveness in handling **large-scale datasets**.

Aria | *OpenAI GPT-4o-mini, LangChain, LaTeX, Python, JSON* | Code

- Developed an intelligent resume generation system leveraging **large language models** (LLMs) to analyze job requirements and dynamically create tailored professional documents.
- Implemented a robust **workflow orchestration** using **LangChain** for multi-stage processing, ensuring high-quality document generation aligned with job postings.
- Achieved a **30% reduction** in time spent on job applications through automated content generation, increasing response rates from recruiters significantly.

Education

University of Manitoba Sep 2023 – Aug 2025
Master of Science in Computer Science (GPA: 4.4 / 4.5) *Winnipeg, Canada*

- **Relevant Coursework:** Security & Privacy, Deep Generative Modeling, Blockchain & Distributed Systems: A+

K.N. Toosi University of Technology Sep 2018 – Feb 2023
Bachelor of Science in Computer Engineering

Technical Skills

AI / Machine Learning: TensorFlow, PyTorch, Scikit-learn, Keras, Transformers
Languages: Python
Cloud & DevOps: Google Cloud Platform, Docker, Git, CI/CD
Databases: PostgreSQL, MongoDB, MySQL
Web Frameworks: *Back-end*: Django, Flask. *Front-end*: React
Tools & Methodologies: Agile, Scrum