

Mohammad Karimi

Senior Data Scientist-KPMG Lighthouse

📍 Edmonton, AB 📞 825-440-5744 ✉ mmdkarimi96@gmail.com **in** [Linkedin](#) [Github](#) [Website](#)

Highlights

- Skilled Senior Data Scientist and ML developer having 6+ years of experience in designing and developing end-to-end ML and data pipelines
- Highly proficient in Natural Language Processing and LLMs, data engineering (ETL), and data science with 6+ years of research and industrial experience in developing large-scale LLMs, data architectures, and machine learning pipelines.
- Demonstrated strong leadership, representation, and teamwork by collaborating with diverse clients, eagerly tackling new challenges and adapting flexibly to evolving project demands
- Achieved advanced cloud certifications including Azure AI-102, DP-203, DP-100, DP-900, AI-900, and **GCP Professional Data Engineer**

Relevant Work Experience

Senior Data Scientist

Jun. 2023 - Present

KPMG Lighthouse, Toronto, ON (Remote)

- Developed a cloud-based solution for Bell Finance on **GCP** using **BigQuery**, **GCS**, **Data Fusion**, and **Cloud Composer** to map customers to billing accounts and enhance revenue reporting
- Developed an AI-driven SOC report **summarizer** using **NLP**, **Transformers**, and local **LLMs** with instruction fine-tuning, achieving over **92%** accuracy in information retrieval
- Contributed to the development of **nine data science pipelines** for **Bell** on **Vertex AI** (Google Cloud Platform), encompassing processes from data ingestion to model deployment
- Transformed and migrated **Prudential Financial's** on-prem solution into **GCP cloud infrastructure**, optimizing **scalability**, security, and cost-efficiency by integrating **chatbots** using **LLMs** to handle quick database **queries**.
- Served as a technical **QA analyst** for **Peoples Trust Bank**, ensuring **test strategy** and **data quality** during the **data and backend migration** from **Peoples Bank of Canada**, applying functional, regression, unit, and system testing.

Associate Machine Learning Developer

Sep. 2022 - May 2023

AltaML, Edmonton, AB

- Developed a bias and hate speech classifier based on **Open AI's GPT 3.5 Turbo API** on **Azure** to classify **hate speech** and different types of **biases** from social media posts
- Developed and optimized a deep learning model pipeline on **Azure** for **City of Calgary** to identify road defects based on Google street view panorama imagery
- Implemented and deployed a quick and **scalable ML pipeline** on **Azure** for **Government of Alberta** to predict on over **2.6 billion** geospatial satellite imagery data points with an accuracy of **+88%** using cross-validation
- Gained a deep understanding of **business values** and **agile delivery model** through close collaboration with the clients and the business team

Data Scientist and NLP Researcher

Jan. 2021 – Aug. 2022

University of Alberta, Edmonton, AB

- Developed and fine-tuned **language models** (i.e., n-grams and recurrent neural networks) to identify grammatically incorrect writing from **big data** resources with an accuracy of **+97%**

- Implemented a data and ML pipeline to process and predict the **price** of airbnb listings using **XGBoost** with a negligible error rate
- Implemented and fine-tuned ML models from scratch to extract the semantic relation between chunks of a sentence with an accuracy of **+87%**

Software Engineer

Sep. 2018 – Jan. 2020

Iranian national robotics team, Tehran, Iran

- Implemented a data acquisition pipeline to collect imagery dataset of objects and employed **YOLO** deep learning algorithm to enable object recognition
- Designed and developed a ML algorithm to enable human robot interaction using speech and prompt processing

Education

M.Sc. in Computer Science

Jan. 2021 – Aug. 2022

University of Alberta

Edmonton, AB

B.Sc. in Software Engineering

Sep. 2015 – Sep. 2019

University of Science and Technology of Mazandaran

Mazandaran, Iran

Technical & Soft Skills

Languages: Python, Java, C++, SQL, C#, JavaScript, PHP, HTML/CSS

NLP: NLTK, Spacy, Gensim, Hugging Face, Stanza, DeepSpeed, PaddlePaddle

Libraries: Tensorflow, Pytorch, Keras, Sklearn, OpenCV, YOLO, Pandas, Numpy

Tools: Azure, Linux, Git, Informatica, Docker, Bash, ROS, Cognos, Google Street View

Database: MySQL, Microsoft SQL Server, SQLite, PostgreSQL

Visualizations: Tableau, Matplotlib, Seaborn, Plotly, Microsoft Power BI, LIME, Streamlit

Soft Skills: Management & Leadership, Critical Thinking, Communication, Work Ethic, Time Management, Creativity

Selected Publications

- **Karimiabdolmaleki, M.**, Farias Wanderley, L., Cutumisu, M., Demmans Epp, C. (2023). Identifying negative language transfer in the English writing of Chinese and Farsi native speakers. European Association for Research on Learning and Instruction (EARLI) Conference, Thessaloniki, Greece, August 22-26.
- **Karimi Abdolmaleki, M.**, Demmans Epp, C., Cutumisu, M. (2022). Automated feedback generation in programming environments: A systematic review. Poster. Systematic Review and Meta-Analysis SIG Poster Session. American Educational Research Association (AERA) Annual Meeting, San Diego, CA, April 22-25.

Community Engagements

ML Instructor

Alberta AI Association

- Teaching an advanced ML course with more than 15 students, covering NLP topics and MLOps, including data processing, model training, model evaluation, and deployment

Guest Speaker

Action for Healthy Communities

- Demonstrated leadership and presentation skills by presenting ML-related topics to people from different backgrounds